Research Topics

Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers, 2nd Edition

Edited by
Connie Qun Guan, Laura M. Morett, Huili Wang
and Wanjin Meng

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-4479-2
DOI 10.3389/978-2-8325-4479-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers, 2nd Edition

Topic editors
Connie Qun Guan — Beijing Language and Culture University, China
Laura M. Morett — University of Alabama, United States
Huili Wang — Dalian University of Technology, China
Wanjin Meng — National Institute For Education Sciences (China), China

Citation
Guan, C. Q., Morett, L. M., Wang, H., Meng, W., eds. (2024). *Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers, 2nd Edition.* Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4479-2

Publisher's note: This is a 2nd edition due to an article retraction.

Table of contents

05 Editorial: Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers
Laura M. Morett

08 Hemispheric Processing of Chinese Scientific Metaphors: Evidence *via* Hemifield Presentation
Min Huang, Lexian Shen, Shuyuan Xu, Yanhong Huang, Shaojuan Huang and Xuemei Tang

19 Gesture–Speech Integration in Typical and Atypical Adolescent Readers
Ru Yao, Connie Qun Guan, Elaine R. Smolen, Brian MacWhinney, Wanjin Meng and Laura M. Morett

30 N400 Indexing the Motion Concept Shared by Music and Words
Tongquan Zhou, Yulu Li, Honglei Liu, Siruo Zhou and Tao Wang

49 Bilingual Processing Mechanisms of Scientific Metaphors and Conventional Metaphors: Evidence *via* a Contrastive Event-Related Potentials Study
Xuemei Tang, Lexian Shen, Peng Yang, Yanhong Huang, Shaojuan Huang, Min Huang and Wei Ren

62 Is the Processing of Chinese Verbal Metaphors Simulated or Abstracted? Evidence From an ERP Study
Ying Li, Xiaoxiao Lu, Yizhen Wang, Hanlin Wang and Yue Wang

72 Facilitative Effects of Embodied English Instruction in Chinese Children
Connie Qun Guan and Wanjin Meng

87 Decoding brain activities of literary metaphor comprehension: An event-related potential and EEG spectral analysis
Lina Sun, Hongjun Chen, Chi Zhang, Fengyu Cong, Xueyan Li and Timo Hämäläinen

102 Metonymy Processing in Chinese: A Linguistic Context-Sensitive Eye-Tracking Preliminary Study
Xianglan Chen, Hulin Ren and XiaoYing Yan

112 Effects of social experience on abstract concepts in semantic priming
Zhao Yao, Yu Chai, Peiying Yang, Rong Zhao and Fei Wang

123 **Language switching may facilitate the processing of negative responses**
Anqi Zang, Manuel de Vega, Yang Fu, Huili Wang and David Beltrán

134 **Effects of language background on executive function: Transfer across task and modality**
Yeonwoo Kim, Zixuan Ye, Zachary Leventhal, Wei-Ju Wang and Erik D. Thiessen

TYPE Editorial
PUBLISHED 01 February 2023
DOI 10.3389/fpsyg.2023.1144176

OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Laura M. Morett
✉ lmorett@ua.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 13 January 2023
ACCEPTED 20 January 2023
PUBLISHED 01 February 2023

CITATION
Morett LM (2023) Editorial: Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers. *Front. Psychol.* 14:1144176. doi: 10.3389/fpsyg.2023.1144176

Editorial: Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers

Laura M. Morett*

Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, Tuscaloosa, AL, United States

KEYWORDS

language embodiment, gesture, handwriting, metaphor, event-related potential (ERP)

Editorial on the Research Topic
Language embodiment: Principles, processes, and theories for learning and teaching practices in typical and atypical readers

This Research Topic endorses the embodied view of language processing, which claims that, rather than being abstract, amodal, and arbitrary, the body plays a central role in shaping it (Glenberg and Kaschak, 2003). In particular, this Research Topic examines the implications of the embodied view for learning and teaching practices for typical and atypical populations, focusing in particular on its cognitive and neural mechanisms.

Several of the papers in this Research Topic focus on metaphor and closely related topics, an important theme in research on language embodiment (Gibbs, 2017; Hampe, 2017). Li et al. helped to disentangle the roles of simulation and abstraction in metaphor processing in Chinese by using event-related potentials (ERPs) to reveal that, in the verb processing stage, verb-object metaphors trigger simulation, whereas subject-verb metaphors trigger abstraction. This finding suggests that verbal metaphors are processed using both simulation and abstraction, and that metaphorical meaning is integrated in real-time during metaphor comprehension. Relatedly, Chen et al. investigated whether processing of metonymy, like metaphor, entails embodied simulation. Using eye-tracking, they found that readers take longer to arrive at a literal interpretation than a metonymic one when preceding information is weakly related to target words but not when it is strongly related to target words. Moreover, they found that preceding and spillover contextual information contributes to metonymy processing when spillover affects metonymy more than literal meaning. These findings provide insight into how sentential components contribute to metonymic processing of target words in Chinese, contributing to a model of metonymy processing.

Metaphor is pervasive in literary and scientific texts, so it is important to consider how it relates to embodied views of language processing within these contexts. Sun et al. used ERPs and event-related spectral perturbations (ERSPs) to investigate the processing of literary metaphors in modern Chinese poetry in comparison to non-literary conventional metaphors and literal expressions outside literary texts. They found more positive P200 and more negative N400 fronto-central components for literary metaphors than for other stimuli as well as increases in the delta and theta frequency bands during different time windows for literary metaphors. Together, these findings indicate that literary metaphor processing has distinct EEG spectral patterns and that it is characterized by early allocation of attention and conscious experience. Similarly, Tang et al. examined comprehension of scientific metaphors in first (L1; Chinese) and second (L2; English) language using ERPs. Relative to L1 Chinese scientific metaphors, L2 English scientific metaphors elicited more negative N400 and less positive late positive components (LPCs) in the parietal region as well as larger late negativities encompassing smaller areas of brain. These findings indicate that non-native and non-dominant language processing entails increased effort, decreased automaticity, and decreased sensitivity to the conventionality of metaphoric meanings. Finally, Huang et al. used hemifield presentation to investigate functional laterality during processing of scientific and conventional metaphors. The results support the fine-coarse coding hypothesis, which posits that the left hemisphere supports integration of the loosely associated domains of scientific metaphor. Moreover, compared to literal word pairs, conventional metaphors elicited higher LPCs during right visual field presentation, whereas scientific metaphors elicited lower LPCs during left visual field presentation. These findings suggest that processing mechanisms differ between novel and conventional metaphors and that right hemisphere plays a special role in novel metaphoric processing during the mapping stage.

Another important theme in this Research Topic is the relationship between domain- general and domain-specific processing. Zang et al. examined how L1–L2 language switching influences negation processing using a story reading and question verification paradigm in which the language was either kept the same or changed across tasks. They found that switching from L1 to L2 facilitated negative compared to affirmative responses, suggesting that domain-general mechanisms of inhibitory control are recruited simultaneously for L1–L2 language switching and negation. Seeking to reconcile conflicting results concerning the effect of multilingualism on executive function, Kim et al. examined the impact of task modality and found that it yields different patterns of performance between monolingual and multilingual participants, suggesting that distance of transfer from everyday language use may provide a viable explanation for these discrepancies. Yao R. et al. used a Stroop-like task to examine gesture-speech integration in deaf and hard of hearing adolescents. The results revealed that deaf and hard of hearing adolescents showed stronger effects of gesture-speech integration as well as semantic congruency than their hearing peers, revealing how sensory experience can influence language processing.

One method that has been instrumental in illuminating the relationship between domain- general and domain-specific processing in this Research Topic is priming, in which a (domain-general) prime influences processing of a (domain-specific) target. Zhou et al. examined the effects of musical primes with and without tempo changes on different classes of words using ERPs. When primed with tempo changes, state verbs and inanimate nouns elicited larger N400 amplitudes than action verbs and animate nouns, suggesting that such priming facilitates processing of action verbs and animate nouns due to the shared concept of motion across music and language. In a conceptually similar vein, Yao Z. et al. used picture-word semantic priming to examine the impact of social experience on processing of social abstract (e.g., friendship and betrayal) and emotional abstract (e.g., happiness and anger) words. Pictures of positive social scenes facilitated processing of positive social abstract words, and pictures of corresponding facial expressions and gestures facilitated processing of positive emotional abstract words, whereas no facilitatory priming was observed for negative pictures and social or emotional abstract words.

These findings provide evidence that social experience selectively affects processing of positive social and emotional language, providing a mechanism for the grounding of these concepts.

In addition to illuminating the mechanisms of embodied language processing, some papers in this Research Topic provide insight into how they can be leveraged to promote language comprehension in instructional settings. Li and Guan examined the effects of hand writing and drawing on development of orthographic representations. Behavioral and ERP data indicated that drawing facilitates visual word recognition in Chinese compared to viewing. Moreover, N170 amplitude correlated positively with N400 amplitude in the drawing condition, suggesting that hand movement may facilitate the neural correlates between early word recognition and later comprehension. Finally, Guan and Meng examined the effect of embodied morphological training by hand writing roots, dragging roots, and gesturing roots on children's L2 word learning. The results revealed that hand writing roots facilitated sound-meaning and word-sound form integration, whereas dragging and gesturing roots facilitated word form-meaning association, providing evidence that embodied morphological training contributes to children's L2 vocabulary acquisition.

In summary, the articles in this Research Topic reveal the inextricable relationship between language processing and embodied experience from several perspectives using cutting-edge behavioral and neuroscience methods. Importantly, they demonstrate how the embodied grounding of language can be leveraged to enhance learning in typically and atypically developing individuals, complementing other recent efforts in this regard (Boieblan, 2022; Hughes-Berheim et al., 2022; Morett et al., 2022). In doing so, they contribute to the growing literature providing insight into the myriad ways in which language processing is embodied and in which the mechanisms of its embodiment can be leveraged to enhance learning in individuals with diverse lived experiences.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Acknowledgments

The author acknowledges her co-editors for their efforts co-editing this Research Topic.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Boieblan, M. (2022). Enhancing English spatial prepositions acquisition among Spanish learners of English as L2 through an embodied approach. *Int. Rev. Appl. Linguist. Lang. Teach*. (In press). doi: 10.1515/iral-2021-0151

Gibbs, R. W. (2017). *Metaphor Wars*. Cambridge: Cambridge University Press.

Glenberg, A. M., and Kaschak, M. P. (2003). "The body's contribution to language," in *Psychology of Learning and Motivation, Vol. 43*, ed B. H. Ross (Cambridge, MA: Academic Press), 93–126.

Hampe, B. (2017). *Metaphor: Embodied Cognition and Discourse*. Cambridge: Cambridge University Press.

Hughes-Berheim, S. S., Cheimariou, S., Shelley-Tremblay, J. F., Doheny, M. M., and Morett, L. M. (2022). Extending gesture's impact on word learning to reading: A self-paced reading study. *Discour. Process.* 59, 646–667. doi: 10.1080/0163853X.2022.2132080

Morett, L. M., Feiler, J. B., and Getz, L. M. (2022). Elucidating the influences of embodiment and conceptual metaphor on lexical and non-speech tone learning. *Cognition* 222, 105014. doi: 10.1016/j.cognition.2022.105014

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 26 May 2022
doi: 10.3389/fpsyg.2022.894715

Hemispheric Processing of Chinese Scientific Metaphors: Evidence *via* Hemifield Presentation

Min Huang[1†], Lexian Shen[1†], Shuyuan Xu[2], Yanhong Huang[1], Shaojuan Huang[1] and Xuemei Tang[1]*

[1] *School of Foreign Studies, Anhui Polytechnic University, Wuhu, China,* [2] *Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an, China*

OPEN ACCESS

Edited by:
Huili Wang,
Dalian University of Technology, China

Reviewed by:
Xiaolu Wang,
Zhejiang University City College, China
Xueyan Li,
Dalian University of Technology, China

***Correspondence:**
Xuemei Tang
tangxuemei@ahpu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: 12 March 2022
Accepted: 27 April 2022
Published: 26 May 2022

Citation:
Huang M, Shen L, Xu S, Huang Y, Huang S and Tang X (2022) Hemispheric Processing of Chinese Scientific Metaphors: Evidence via Hemifield Presentation. Front. Psychol. 13:894715. doi: 10.3389/fpsyg.2022.894715

The role of the two hemispheres in processing metaphoric language is controversial. In order to complement current debates, the current divided visual field (DVF) study introduced scientific metaphors as novel metaphors, presenting orientation mapping from the specific and familiar domains to the abstract and unfamiliar domains, to examine hemispheric asymmetry in metaphoric processing. Twenty-four Chinese native speakers from science disciplines took part in the experiment. The participants were presented with four types of Chinese word pairs: scientific metaphors, conventional metaphors, literal word pairs, and unrelated word pairs. The first word in each pair was presented centrally, and the second was presented to the left visual field (the Right Hemisphere) or the right visual field (the Left Hemisphere). Event-related potentials (ERPs) were recorded when participants read the target words and judged whether words in each pair were related. The data demonstrated that both hemispheres were involved at the initial stage of metaphor processing, but the right hemisphere took a more privileged role. The significant activation of the left hemisphere for scientific metaphoric processing supports the fine-coarse coding hypothesis. During right-visual-field presentation, the left hemisphere, responsible for the processing of closely related domains, has to integrate the loosely associated domains of scientific metaphor, which greatly increased cognitive taxes. Moreover, the data of late positive components (LPCs) revealed different hemispheric activation between scientific metaphors and conventional metaphors. Compared with literal pairs, conventional metaphors elicited significantly higher LPCs during right visual field presentation, while the scientific metaphor elicited significantly lower LPCs during left visual field presentation. These results suggest different processing mechanisms between novel metaphors and conventional metaphors and the special role of the right hemisphere in novel metaphoric processing at the later mapping stage.

Keywords: scientific metaphor, divided visual field diagram, hemispheric asymmetry, N400, LPC

INTRODUCTION

Metaphors are used pervasively in our daily life to enable more effective and efficient communication. As in "*Marriage is gambling*," two conceptual areas are equated, and readers can easily go beyond the literal meaning to grasp the figurative meaning by analogizing or comparing the two concepts. In previous research works, metaphors were categorized into conventional metaphors and novel metaphors according to their familiarity or conventionality, determining those which can be comprehended effortlessly by the native speakers as conventional metaphors. To date, much attention has been paid to the hemispheric processing of metaphors. However, there is still no agreement, especially on the role of the right hemisphere in novel metaphoric processing.

To complement the current debate, we attempted to further differentiate novel metaphors. The scientific metaphors we used in this study are rather imaginative and often used by scientists to reason or communicate abstract scientific concepts. For example, through the scientific metaphor, "*sound wave*," we may understand the characteristics of "*sound*" comparatively more easily through the concept of "*wave*." Similar to the novel metaphors used in most previous research, the scientific metaphors we used in our study were non-conventional and unfamiliar. However, they have a "more complicated contextual structure" involving a longer mapping distance (Tang et al., 2017a). The source domain of wave (as in "*water wave*") is derived from daily-use contexts, while the target domain in which the metaphor is applied as "*sound*" is from scientific contexts, which increases the difficulty of integration. Besides, the late processing of scientific metaphors involves a unique reasoning process to understand the related knowledge, which taxes semantic reintegration. Using scientific metaphors to study hemispheric processing might shed some light on relevant studies.

Lateralization in Metaphor Processing

Most neuro-linguists and psychologists agree that metaphorical meaning processing is different from literal meaning processing with the former showing a preference for the right hemisphere. Compared with the theory of the left hemisphere involvement in metaphorical processing, *the right hemisphere engagement theory* does gain more attraction, because it fits into a more holistic picture of the brain's division as described in Jung-Beeman's (2005) fine-coarse hypothesis: the left hemisphere is primarily responsible for the fine coding of closely related meanings, while the right hemisphere is for the coarse coding of non-literal meaning, including metaphors.

Moreover, some clinical studies involving schizophrenia, Asperger's syndrome, Alzheimer's disease, and other pathologies have reported irregular lateralization of metaphor processing, indicating the important role of the right hemisphere in metaphor processing (Rapp et al., 2004; Lauro et al., 2008; Ianni et al., 2014). In addition, some experimental studies on healthy subjects also confirmed the right-hemisphere advantage in metaphor processing. For example, a positron emission tomography study found that the processing of metaphorical sentences resulted in increased blood flow in the prefrontal cortex, middle temporal gyrus, anterior cingulate gyrus, and posterior cingulate gyrus of the right hemisphere, compared with the processing of the literal meaning of the same structure (Bohrn et al., 2012). Some fMRI (Yang, 2014; Lai et al., 2015) and Event-related potential (ERP) studies (Tang et al., 2017a,b) also suggested that the right brain played an important role in metaphor processing.

Different linguists have given different explanations for the special role of the right hemisphere in metaphor processing. According to the Fine Coarse Semantic Coding Model (Beeman, 1998, 2005), the right hemisphere has an advantage in semantic processing with a large semantic span, while the left hemisphere is mainly responsible for processing conventional concepts of words. However, based on the graded salience hypothesis (GSH) (Giora, 1997, 2003), the right hemisphere has an advantage in low salient meaning processing, while the left hemisphere has an advantage in high prominent meaning processing (Mashal and Faust, 2008).

In contrast to these findings, other studies reported no predominance of the right hemisphere during non-literal language processing. Some experiments have found that the left hemisphere advantage was found in figurative language comprehension (Eviatar and Just, 2006; Bohrn et al., 2012; Mitchell et al., 2016), while other studies reported that both hemispheres are involved in metaphoric meaning processing (Yang et al., 2016).

Previous studies have found that the salience, conventionality, familiarity, anticipation, predictability, or transparency of metaphors and other figurative languages used in those experiments can modulate the results produced to a large extent. For example, a study on idioms argued that the predictability of metaphors was the main factor that decided the means of semantic processing (Sela et al., 2017). In addition, according to *the metaphor career theory* (Bowdle and Gentner, 2005), metaphors with different degrees of conventionality have different neural mechanisms for processing. According to *the structure mapping theory* (Wolff and Gentner, 2011), metaphor processing involves an initial processing stage of symmetric alignment and a later directional phase in which inferences are projected to the target. More importantly, Wolff and Gentner (2011) proposed that the base of highly conventional metaphors already possesses a salient conventional metaphoric meaning, whereas metaphoric abstraction must be derived anew for a novel figurative. However, some early experiments only distinguished literal meaning from metaphorical meaning without a further classification of metaphors (Eviatar and Just, 2006; Bohrn et al., 2012; Mitchell et al., 2016). Most current studies differentiated conventional metaphors from novel metaphors without further classifying novel metaphors. One of our previous studies adopted a central visual presentation paradigm to compare Chinese poetic and scientific metaphors in sentence context and reported larger late negativity in the LPC window simultaneously on the left and the right hemispheres suggesting both hemispheres of the brain work together when processing scientific metaphors (Tang et al., 2017b). There is an urgent need to study some specific types of novel metaphors to observe the lateralization in metaphor processing.

Hemifield Priming and Metaphor Processing

Many previous research works adopted a divided visual field (DVF) priming paradigm to study the hemispheric effect of metaphor processing. Presenting stimuli in the right or left visual field can control information selectively activating only the left or right visual cortex during the initial stage of language processing. In other words, for normal individuals, information is rapidly transmitted to the opposite hemisphere.

Meanwhile, in traditional behavioral DVF experiments, researchers often modulated stimulus onset asynchrony (SOA) to study lateralization as well as the dynamic time course during metaphor comprehension. Anaki et al. (1998) found that when SOA was short (200 ms), both literal and metaphorical pairs produced a left-brain priming effect, but only metaphorical words produced a right-brain priming effect. When the SOA was long enough (800 ms), only literal words had left-brain priming effects, while metaphorical words had right-brain priming effects. These findings suggested that metaphorical meaning is initially activated in both hemispheres, but the degree of activation declines rapidly in the left hemisphere but remains constant in the right hemisphere. Anaki's experiment did not distinguish the semantic salience of the metaphorical corpus. Faust and Mashal (2007) divided the metaphor sentences into novel metaphors and conventional metaphors in their study. The study found that, under the conditions of SOA of 400 and 1,100 ms, both literal meaning and metaphorical meaning showed priming effects in the left and the right hemispheres, which indicated that both hemispheres worked together during metaphor comprehension in different stages. However, novel metaphors were processed faster and more accurately in the right hemisphere, supporting the right hemisphere theory. The rationale they put forward was that the left brain used sentence constraints to select and integrate literal and metaphorical meanings related to the context of the sentence. The right brain might be less sensitive to sentence context and only participate in alternative interpretations in cases where the literal meaning cannot be explained. However, using the same DVF paradigm, Forgács et al. (2014) found that novel metaphor pairs, like literal and conventional word pairs, have no significant right hemisphere effect and are processed faster and more accurately in the left hemisphere.

In addition, the DVF paradigm was sometimes combined with human electrophysiology to study non-literal language processing. A study reported that jokes presented to the right visual field-left hemisphere (rvf-LH) elicited a larger N400 than the non-joke endings; however, when presented to the left visual field-right hemisphere (lvf-RH), the joke and non-joke endings elicited N400s of equal amplitude (Coulson et al., 2005), supporting the idea that the right hemisphere plays a special role. Another hemifield study on metaphor processing, on the contrary, found that ERP metaphoricity effects were very similar across hemifields, suggesting that the integration of metaphoric meanings was similarly taxing for the two hemispheres (Coulson and Van Petten, 2007). A DVF fMRI study on Chinese idioms reported similar results (Yang et al., 2016).

The Current Study

In order to address the current debates on lateralization during metaphor processing, the present study chose Chinese scientific metaphors to assess the hemispheric processing of novel metaphors *via* a combination of the DVF paradigm and ERP methodology. Both source and target domains of conventional metaphors come from daily life, presenting symmetrically analogical mapping in which a structural alignment between two represented situations is established. For example, "*月亮 (crescent moon)*" and "*小船 (boat)*" have similar shapes and their associated abstract schemata are sufficiently accessible. However, source domains in scientific metaphors are abstract scientific concepts, the processing of which often needs to activate specific concepts in target domains. When understanding the term "*电流 (electric current)*," it is hard to consider it as a pure category name. It involves a comparison between "*电子 (electronics)*" and "*水流 (current)*," presenting orientation mapping from the specific and familiar domains to the abstract and unfamiliar domains. We compare "*human eyes*" to "*cameras*" as they work in a similar way.

When considering the notions of a "*sound wave*" or an "*electric current*," the everyday concepts "*wave*" and "*current*" can help readers quickly understand certain scientific concepts "*sound*" or "*electricity*." Such mapping across daily and scientific concepts requires semantic integration based on the adequate comparison, which results in increased cognitive load in retrieving stored conceptual knowledge and in integrating seemingly unrelated information from different domains in the process of metaphoric comprehension. Our previous research also proved that scientific metaphors elicited a diverse LPC from novel poetic metaphors and conventional metaphors (Tang et al., 2017b).

Despite limitations on spatial resolution, ERP technology has excellent performance on temporal revolution. Compared with other methodology, ERPs provide a continuous measure of word processing that is sensitive to lateralized brain activity over the different stages of processing. The combination of ERP and DVF techniques can help study lateralization during metaphor processing in a more accurate and dynamic way.

Our study replicated Coulson and Van Petten's (2007) research, but used different stimuli and contexts. We used word pairs instead of sentences to avoid the influence of sentence contexts. Moreover, we replaced the novel metaphors with scientific metaphors as their source domain and target domain are from different contexts and therefore have a longer mapping distance. According to structure mapping theory, we assumed that scientific metaphors would elicit more negative N400s or other late negativity components. More importantly, we supposed that the presentation side of stimuli would interact with the ERP effects of metaphoric variables. If hemispheric differences in semantic activation affect metaphor comprehension, the presentation side (visual field) would be expected either to facilitate processing or to make it more difficult.

In the previous ERP studies, the amplitudes of N400 were reported to indicate the degree of difficulties in retrieving and integrating contextual meaning in metaphor comprehension (Lai and Curran, 2013; Schneider et al., 2014). Metaphors should thereby elicit a more negative N400 than literal meaning. In addition, LPC as a late component was reported to reflect a secondary integration of meaning required by novel metaphors

TABLE 1 | Chinese sample stimuli.

	Word 1	Word 2	English meaning of Word 1	English meaning of Word 2
Scientific metaphors	淋巴	警察	Lymph	Police
SM	电子	行星	Electrons	Planets
	导体	隧道	Conductors	Tunnels
	染色体	姐妹	Chromosomes	Sisters
	病毒	杀手	Virus	Killer
Conventional	语言	桥梁	Language	Bridge
metaphors	杭州	天堂	Hangzhou	Heaven
CM	家庭	港湾	Home	Harbor
	手机	伙伴	Cellphone	Partner
	恋爱	咖啡	Love	Coffee
Literal expressions	教授	学者	Professor	Scholar
LT	汉语	语言	Chinese	Language
	北京	城市	London	City
	蚂蚁	昆虫	Ant	Insect
	小狗	宠物	Dog	Pet

(Kazmerski et al., 2003; Rutter et al., 2012; Tang et al., 2017a,b). Therefore, if the right hemisphere takes a special role in metaphor comprehension, our hypotheses are as follows.

(1) The amplitude of N400s elicited by scientific metaphors should be larger than those of conventional metaphors and literal pairs with the presentation to both hemifields.
(2) There should be some interaction between conditions and sides. Specifically, the differences between N400s and LPCs elicited by scientific metaphors, conventional metaphors, and literal pairs should be more significant with the presentation to rvf-LH than lvf-RH.

EXPERIMENTAL PROCEDURE

Participants

Twenty-four undergraduate students (14 males, 10 females, average age 21.5) from Anhui Polytechnic University, aged 18–22, participated in the ERP experiment. All the participants were from science disciplines considering the possible difficulties in understanding the academic knowledge involved in scientific metaphors. All were native Chinese speakers and yielded a laterality quotient of at least +80 on the Edinburgh Inventory, indicating right-handedness (Oldfield, 1971). Exclusion criteria were sinistrality, past or present psychiatric illness or neurological disorder, or major head injury. All participants had normal or corrected to normal vision and were given monetary compensation for their participation. The experimental standards of the study were approved by the local Review Board for Human Participant Research. Each subject provided written informed consent before participating. They were presented with four types of word pairs and asked to perform semantic judgment on the second word of each pair that was presented to the rvf-LH or lvf-RH. Four participants had to be excluded from data analyses due to low accuracy in semantic judgment tasks, resulting in a final sample size of 20 subjects (12 males, 8 females, average age of 21.55).

Stimuli

The stimuli of the experiment were better calibrated based on the corpus of ERP experimental studies on the neural mechanism of Chinese metaphor understanding (Tang et al., 2017a). The stimulus pool consisted of 375 pairs of words, all in Chinese. They were grouped into four types of semantic relations: literal pairs (LT: *debris flow* 泥石流, *disaster* 灾害), conventional metaphoric (CM: *eyes* 眼睛, *window* 窗口), scientific metaphoric (SM: *electric current* 电流, *current* 水流), or unrelated pairs (UR: *balcony* 阳台, *Antarctic Circle* 南极圈) (see **Table 1** for examples).

The first word of each pair served as the prime and the second as the target. Each word contained two or three Chinese characters. Besides, there were a certain number of scientific words in both literal and unrelated word pairs and all the scientific terms or concepts were collected from middle school or high school textbooks, so as to achieve a balance between different conditions in terms of word frequency.

Prior to the neurophysiological study, several pretests were conducted by raters who did not participate in the ERP experiment. Firstly, to determine the degree of semantic relatedness for the word pairs in each condition, 60 raters were presented with a list containing all 375-word pairs

TABLE 2 | The results of pretests.

	Meaningfulness		Figurativeness		Familiarity	
	M	SD	*M*	SD	*M*	SD
SM	3.5	0.6	3.3	0.24	2.85	0.59
CM	3.8	0.42	3.79	0.24	3.64	0.45
LT	4.09	0.44	1.4	0.19	4.12	0.24
UR	1.45	0.2			1.43	16.24

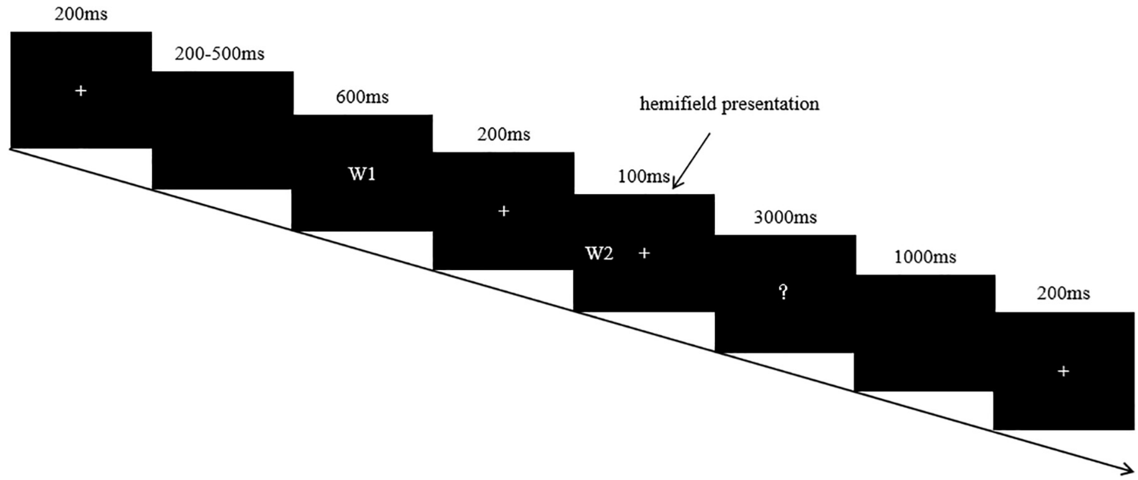

FIGURE 1 | Experimental paradigm. (W1 and W2 refer to word 1 and word 2 respectively).

(75 scientific metaphoric word pairs, 75 conventional, metaphoric word pairs, 75 literal word pairs, and 150 unrelated word pairs) and asked to determine the plausibility and familiarity of all the word pair on a 1–5 scale (1 = not plausible/familiar, 5 = extremely congruent/familiar). Then another 60 raters were asked to decide the figurativeness of the scientific metaphoric, conventional metaphoric, and literal word pairs on a 1–5 scale (1 = not figurative, 5 = extremely figurative).

According to the results of the plausibility and familiarity judgments (see **Table 2** for the descriptive data of pretests), 40 pairs with familiarity and plausibility over 3.5 (average rating of 4.1) and figurativeness below 1.7 (average rating of 1.4) were selected as the literal pairs, and 60 pairs with familiarity and plausibility below 1.7 (average rating of 1.4) were selected as the filler pairs. Among the 75 pairs of scientific metaphors, pairs with figurativeness lower than 2.5 were removed, and the remaining pairs were selected again according to the degree of familiarity and plausibility. Finally, 40 pairs with an average familiarity of over 2.9, plausibility over 3.5, and figurativeness over 3.3 were chosen as scientific metaphors. In 75 pairs of daily metaphors, pairs with figurativeness lower than 3 were eliminated, and then the remaining pairs were further evaluated according to their plausibility and familiarity. Finally, 40 pairs with an average figurativeness (see **Table 2**), plausibility and familiarity rating of 3.8, 3.8, and 3.6 were selected as daily metaphorical word pairs respectively. A repeated-measures ANOVA yielded significant differences ($ps < 0.01$) between word pair types in all dimensions.

Procedures

The experiment took place in a sound-attenuated, electrically shielded room. Before the experiment, considering the possible difficulties in understanding the academic knowledge related to scientific words, we let participants read a list of the relevant terms alongside their definitions. Participants also had the opportunity to look up any unfamiliar ones using their cell phones or consult us to verify meanings. Participants were required to put their jaws on a stent fixed on a small table. The distance between the eyes of the participants and the display screen was 60 cm. The participants were asked to judge whether the priming words (*电荷/electronics*) and the target words (*水流/current*) were semantically related. Semantic relevance was defined as some similarities in appearance, nature, function, and working principle between priming words and target words, and irrelevance was defined as no similarities. E-Prime 2.2 was used to edit and present stimuli. Four types of word pairs were presented pseudo-randomly. Each word was presented in white on a black background. In terms of the font used, we opted for a regular script at a font size of 50. The first word of each pair was presented in the central field of vision as a priming word, and then a target word was presented in the left or right field of vision, with an average angle of 1.9° to the fixation in the central field. ERPs were recorded when participants read the target words and indicated whether words in each pair were semantically related. Behavioral data of semantic judgment was also recorded.

Stimuli on each trial were presented in the following time sequences: fixation cross (200 ms), blank (200–500 ms), priming word (600 ms), fixation cross (200 ms), target word (100 ms), and question mark (3,000 ms). Our reason for selecting an 800 ms SOA was that the metaphoric meaning is more adequately processed under this condition according to previous research. At the sight of the question mark, participants gave their judgment on whether the two words in each pair were semantically related or not. This was done by pressing a corresponding key as quickly/accurately as possible, using their left or right index fingers. An upper time limit of 3,000 ms was permitted for responses and was followed by a 1,000 ms inter-trial interval. The overall sequence of events for a trial is illustrated in **Figure 1**. Before the main session of the experiment, there was a brief practice session to familiarize the participants with the experimental procedure.

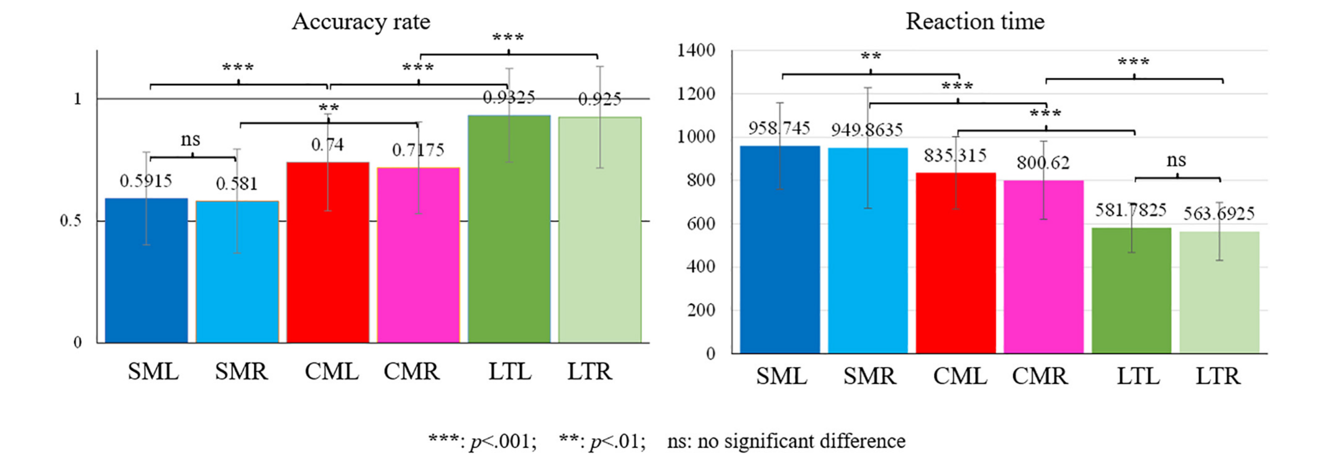

FIGURE 2 | Behavioral results. (The third letter after stimulus type refers to the visual field. For example, SML refers to scientific metaphors presented to left visual field).

EEG Recording and Analysis

EEG readings were continuously recorded from 64 scalp sites at a sampling rate of 512 Hz. Electrode impedance was kept below 5 kΩ. EEG epochs were synchronized with the onset of stimulus presentation and analyzed by all-caps MATLAB. Computerized artifact rejection was performed before averaging to discard epochs in which eye movements, blinks, excessive muscle potentials, or amplifier blocking occurred. EEG epochs associated with an incorrect behavioral response were also excluded. The artifact rejection criterion was a peak-to-peak amplitude exceeding 50 μV. This resulted in a rejection rate of ∼5%. ERPs were averaged off-line from −200 ms before stimulus onset to 1,000 ms after. ERP components were identified and measured, with reference to the average baseline voltage over the interval from −100 to 0 ms, at sites and latency where they reached their maximum amplitude.

RESULTS

Behavioral Results

A 3 condition (literal, conventional, scientific) × 2 view (lvf, rvf) two-way ANOVAs yielded significant main effects of condition for reaction time [$F(2,38) = 62.95$, $p = 0.000$, $\eta_p^2 = 0.768$]. Pairwise comparison showed that the reaction time of scientific metaphors was significantly longer than that of conventional metaphors and literal pairs (*ps* = 0.000), while the reaction time of conventional metaphors was also significantly longer than that of literal pairs ($p = 0.000$). For accuracy rate, a main effect between conditions was found [$F(2,38) = 39.985$, $p = 0.000$, $\eta_p^2 = 0.678$]. Pairwise comparison showed that the accuracy rate of scientific metaphors was significantly lower than that of conventional metaphors and literal pairs (*ps* = 0.000), while the accuracy rate of conventional metaphors was significantly lower than that of literal pairs ($p = 0.000$). No distinguished hemifield presentation effect, as well as interactions with the condition, was found for reaction time and accuracy rate. Statistical results were shown in **Figure 2**.

TABLE 3 | Electrodes chosen for data analysis.

	Left	Midline	Right
Frontal	F7, F3, FT7, CP3	Fz	F4, F8, FT4, FC8
Central	T7, C3	Cz	C4, T8
Parietal	TP7, CP3, P7, P3	CPz, Pz	TP4, CP8, P4, P8

Electrophysiological Results

According to the previous research, we chose Fz, Cz, CPz, Pz as midline, F3, FC3, C3, CP3, P3, F4, FC4, C4, CP4, P4 as dorsal area, F7, FT7, T7, TP7, P7, F8, FT8, T8, TP8, P8 as ventral area, and F7, F3, Fz, F4, F8, FT7, FC3, FC4, FT8 as an anterior area, T7, C3, Cz, C4, T8 as a central line, and TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P2, P4, P8 as a posterior area in order to further analyze the hemisphere processing in metaphors (as shown in **Table 3**).

The resulting amplitudes of N400 and LPC were entered into 3 condition × 2 view (rvf, lvf) × 3 region (frontal, central, parietal) × 3 brain area (left, midline, right) four-way ANOVAs for repeated measures.

100–200 ms

In the time window of N1 (100–200 ms), consistent with **Figure 3**, a repeated-measures ANOVA showed a significant view × hemisphere interaction [$F(2,38) = 25.337$, $p = 0.000$, $\eta_p^2 = 0.571$]. *Post hoc* analysis showed that in the left hemisphere, stimuli presented to rvf-LH elicited significantly higher N1 than stimuli presented to lvf-RH [$F(1,19) = 12.674$, $p = 0.002$, $\eta_p^2 = 0.400$], while in the right hemisphere, stimuli presented to lvf-RH elicited significantly higher N1 than stimuli presented to rvf-LH [$F(1,19) = 15.433$, $p = 0.001$, $\eta_p^2 = 0.448$].

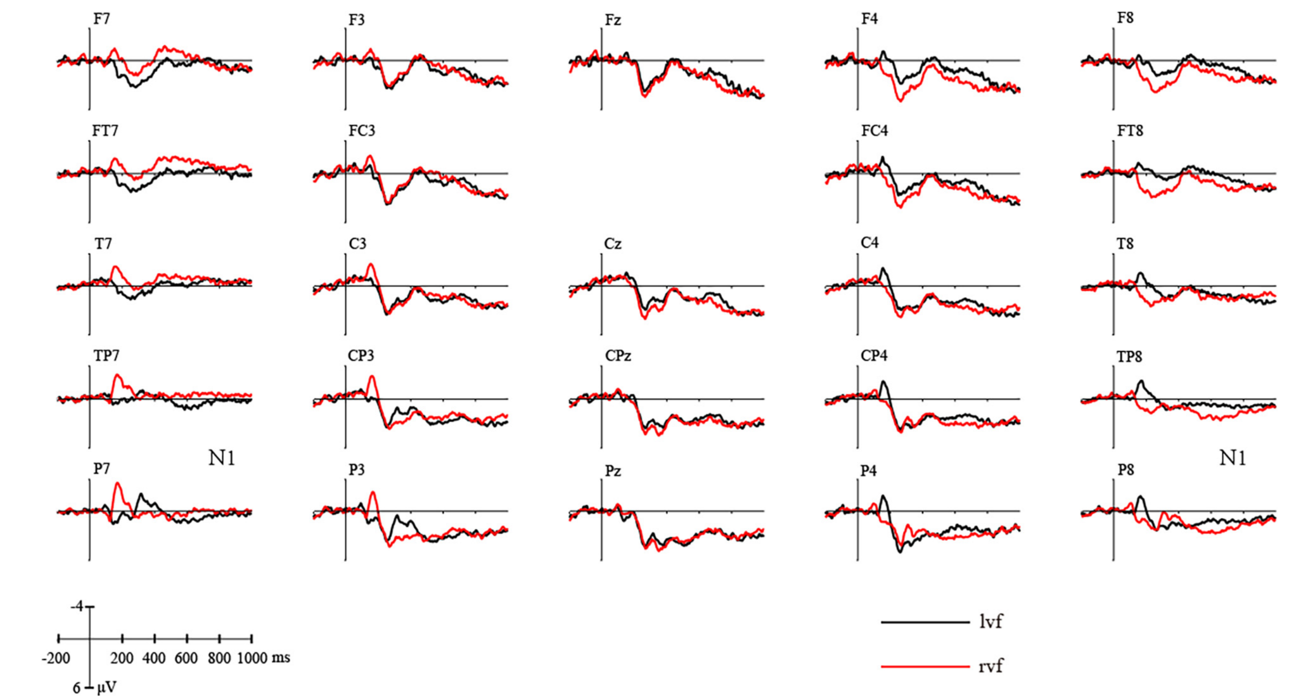

FIGURE 3 | Difference curves of the grand average ERPs elicited by stimuli presented on left visual field and right visual field.

N1/P1 was assumed to be related to spatial attention, enhanced by alterations to the spatial factor (Omoto et al., 2001). The clear negative response observed around 180 ms after stimuli onset underlined the validity of our experiments.

370–550 ms

In the N400 time window (370–550 ms), the condition × side (rvf-LH presentation, lvf-RH presentation) × region (frontal, central, and parietal) ANOVA revealed a significant main effect of condition [$F(2,38) = 10.266$, $p = 0.001$, $\eta_p^2 = 0.351$]. Scientific metaphors elicited more negative N400 ($M = 0.260$, SD = 0.469) than conventional metaphors ($M = 1.613$, SD = 0.557) and literals ($M = 2.329$, SD = 0.701). *Post hoc* analysis showed significant difference between N400s elicited by scientific metaphors and literal pairs ($p = 0.004$) as well as scientific metaphors and conventional metaphors ($p = 0.010$), while no significant difference was found between N400s elicited by conventional metaphors and literal metaphors ($p = 0.304$). There was no significant difference between sides [$F(1,19) = 0.732$, $p = 0.403$, $\eta_p^2 = 0.037$] but a marginally significant interaction between condition and side [$F(2,38) = 2.650$, $p = 0.099$, $\eta_p^2 = 0.122$]. *Post hoc* analysis showed differences of N400s elicited by scientific metaphors, conventional metaphors and literal pairs were more significant during rvf-LH presentation ($p = 0.008$) than lvf-RH presentation ($p = 0.015$).

Further ANOVA for scientific metaphors and literal pairs revealed no significant condition × side interaction ($p = 0.353$). The differences between N400s elicited by scientific metaphors and literal pairs were significant in both visual field presentations. Moreover, the ANOVA for conventional metaphors and literal pairs showed that conventional metaphors elicited significantly higher N400s than literal pairs during lvf-RH presentation ($p = 0.050$) as opposed to rvf-LH presentation ($p = 0.581$).

More interestingly, the ANOVA for scientific metaphors and conventional metaphors revealed a significant main effect of condition [$F(1,19) = 11.359$, $p = 0.003$, $\eta_p^2 = 0.374$] and also a marginally significant interaction between condition and side [$F(1,19) = 3.369$, $p = 0.082$, $\eta_p^2 = 0.151$]. *Post hoc* analysis showed no significant difference between N400s elicited by scientific metaphors ($M = 0.808$, SD = 0.490) and conventional metaphors ($M = 1.454$, SD = 0.615) during lvf-RH presentation ($p = 0.229$), but a significant difference between N400s elicited by scientific metaphors ($M = -0.288$, SD = 0.574) and conventional metaphors ($M = 1.773$, SD = 0.768) during rvf-LH presentation ($p = 0.002$) (as shown in **Figures 4**, **5**).

700–900 ms

In the LPC time window (700–900 ms), the condition × side × region ANOVA revealed a significant main effect of condition [$F(2,38) = 5.242$, $p = 0.019$, $\eta_p^2 = 0.216$] with no interactions. *Post hoc* analysis showed marginally significant difference ($p = 0.074$) between LPCs elicited by scientific metaphors and literal pairs ($M = 3.095$, SD = 0.558) and marginally significant difference ($p = 0.062$) between those by scientific metaphors ($M = 1.706$, SD = 0.429) and conventional metaphors ($M = 2.878$, SD = 0.435) and no significant difference between those by conventional metaphors and literal pairs ($p = 1.000$).

Similar to the results of N400, ANOVA for scientific metaphors and literal pairs revealed a significant main effect of

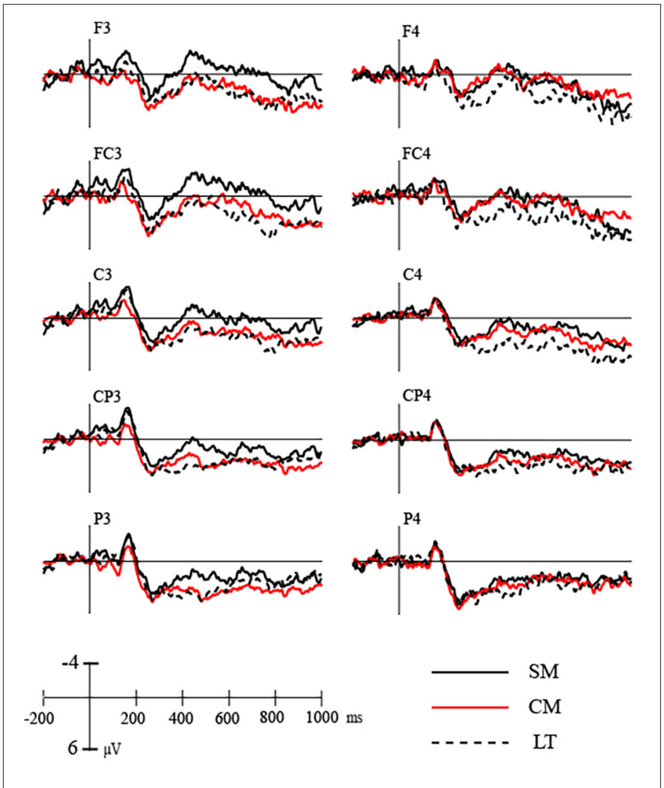

FIGURE 4 | Grand average ERPs elicited by stimuli during left hemisphere presentation (F3–P3) and right hemisphere presentation (F4–P4).

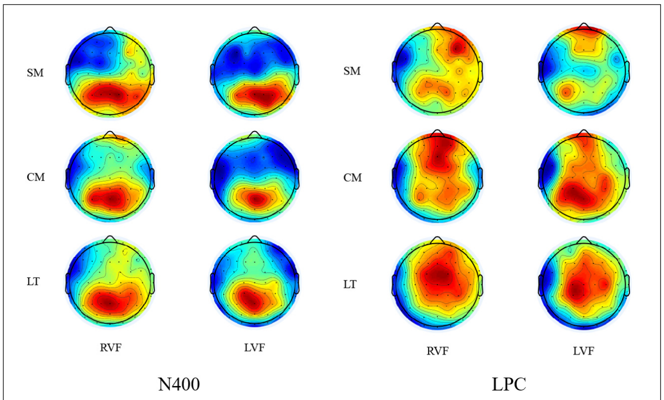

FIGURE 5 | Topographic plots of stimuli for the N400 (370–550 ms) window and the LPC (700–900 ms) window.

condition [$F(1,19) = 5.945$, $p = 0.025$, $\eta_p^2 = 0.238$]. Scientific metaphors elicited a significantly lower LPC than literal pairs during lvf-RH presentation ($p = 0.017$) but not during rvf-LH presentation ($p = 0.177$). For conventional metaphors and literal pairs, there was a marginally significant condition × side interaction [$F(1,19) = 4.354$, $p = 0.051$, $\eta_p^2 = 0.186$] with conventional metaphors eliciting distinguished higher LPC only during rvf-LH presentation ($p = 0.084$).

Moreover, the ANOVA for scientific metaphors and conventional metaphors revealed a significant main effect of condition [$F(1,19) = 6.356$, $p = 0.021$, $\eta_p^2 = 0.251$]. Scientific metaphors elicited significantly lower LPC than conventional metaphors during rvf-LH presentation ($p = 0.048$) but not during lvf-RH presentation ($p = 0.145$). See **Table 4** below.

DISCUSSION

The current DVF study introduced scientific metaphors as novel metaphors to examine hemispheric asymmetry. The behavioral results showed increased reaction time and decreased accuracy rate for scientific metaphor comprehension compared to the processing of conventional metaphors and literal pairs, indicating a special mechanism of novel metaphoric processing. But due to limitations of behavioral data, it failed to show any significant differences between rvf-LH and lvf-RH presentation.

The use of ERP in our study clearly revealed the role of each hemisphere in metaphor processing in a temporally dynamic way. Our study reported the same reversing asymmetries of the visual N1 on either hemifield as in Coulson and Van Petten's (2007) study. In the right hemisphere, stimuli with lvf-RH presentation elicited a larger N100 compared to stimuli with rvf-LH presentation, while in the left hemisphere, stimuli with rvf-LH presentation elicited a larger N100 compared to that with lvf-RH presentation. Although the effect was insensitive to word pair types, it confirmed that lateral processing is evident when stimuli were presented to the opposite visual field, indicating the validity of the study at least in the early stage. In the N400 time window, the results showed that scientific metaphors elicited higher N400 than literal pairs during both rvf-LH and lvf-RH presentation while conventional metaphors elicited higher N400 than literal pairs only during lvf-RH presentation. More interestingly, scientific metaphors elicited higher N400 than conventional metaphors with an interaction between conditions and sides showing a significant difference only during rvf-LH presentation. In the LPC time window, the data suggest that both hemispheres are involved when processing figurative languages, but the right hemisphere takes a more privileged role. The significant activation of the left hemisphere for scientific metaphor processing supports the fine-coarse coding hypothesis (Beeman, 2005). During rvf-LH presentation, the left hemisphere, responsible for the processing of closely related domains, has to integrate the loosely associated domains of scientific metaphor, which greatly increases cognitive taxes.

N400 (370–550 ms)

Consistent with our prediction, scientific metaphors elicited higher a N400 than conventional metaphors and literal pairs in both hemispheres, indicating a unique mechanism of processing novel metaphors. The result is consistent with some of our previous studies with a central presentation in eliciting larger N400s for both scientific metaphors and conventional metaphors as compared to literal pairs in sentence context (Tang et al., 2017a).

With presentation to rvf-LH, scientific metaphors elicited higher N400 than both conventional metaphors and literal pairs

TABLE 4 | *Post hoc* analysis results of LPCs elicited by the three conditions.

	LH presentation		RH presentation	
SM	*M* = 1.563, SD = 0.522		*M* = 1.849, SD = 0.471	
		$p = 0.048$		$p = 0.145$
CM	*M* = 2.969, SD = 0.584		*M* = 2.787, SD = 0.497	
		$p = 0.084$		$p = 0.118$
LT	*M* = 2.380, SD = 0.572		*M* = 3.811, SD = 0.741	

supporting the right hemisphere advantage in novel metaphoric processing in the early stage. According to Aurnhammer et al. (2021), lexical retrieval is indexed by the N400, which is sensitive to linguistic properties like frequency, association, and expectancy. Compared with conventional metaphors and literal pairs, the primary words of scientific metaphors are usually scientific terms such as "*electronics*" and hereby have lower frequency as well as weaker association with base word "*current*" than conventional metaphors or literal pairs. This reveals the unique semantic structure of scientific metaphors that involves two different contexts (Tang et al., 2017a), which may cause greater cognitive taxes for lexical retrieval at the left hemisphere.

During lvf-RH presentation, scientific metaphors and conventional metaphors also elicited higher N400s than literal pairs indicating the unique processing of figurative language in the right hemisphere. According to *the structure-mapping theory* (Wolff and Gentner, 2011), processing a metaphor involves an initial alignment between target and base domains. Therefore, the processing of literal pairs may only involve information retrieval and mainly activate the left hemisphere, while the processing of metaphor may involve not only information retrieval but also the structure alignment, which may activate the right hemisphere to compare the target and base to find correspondences at this stage.

In addition, conventional metaphors elicited higher N400 than literal pairs only in the right hemisphere. This finding proves that conventional metaphors have a different processing mechanism from literal meaning processing: the right hemisphere is involved in semantic integration. And more importantly, there was no significant difference between N400s elicited by scientific metaphors and conventional metaphors during lvf-RH presentation, but a significant difference between N400s elicited by scientific metaphors and conventional metaphors during rvf-LH presentation, which reveals a special role of the right hemisphere, especially in novel metaphor comprehension. This is to say, compared to the left hemisphere, the right hemisphere is more sensitive to the semantic integration between conceptual domains over a long distance.

Our findings, however, are not aligned with Coulson and Van Petten's (2007) study in which ERP metaphoricity effects were very similar across hemispheres despite the left-hemisphere advantage in processing low-cloze literals. One possible reason for the contradictory results lies in the different stimuli used in the two studies. In their research, the concepts of both source domain and target domain of the low-cloze metaphors were mostly from everyday contexts. Firstly, the metaphoricity of those novel metaphors might not have been high enough to elicit significantly different ERP components. Scientific metaphors are of different mapping distances and contextual complexity, which can amplify processing differences between metaphors and literal meaning. Secondly, the individual differences of the participants in terms of their general knowledge might also have brought some impact on the experimental results. For example, a bibliophile might have a totally different processing mechanic from a bibliophobe when processing poetic metaphors. Those individual differences might not be able to be manifested through education or age, and ignorance of these types of differentials might lead to inaccuracy. However, in our study, the scientific metaphors we used as the experimental corpus were obtained from common scientific concepts which are taught in middle school or high school in China. More importantly, there is a learning process before the formal experiment, which can effectively reduce the processing differences caused by different general contextual knowledge of the participants. Besides, the sentence context used in the experiments might be another reason why no significant differences in N400 elicited by low-cloze metaphors compared to literal meaning in the study. The left hemisphere advantage has been found in integrating sentence contexts (Forgács et al., 2014), which might eliminate the differences caused by the conditions.

Late Positive Component (700–900 ms)

Unlike in previous DVF experiments, our study also reported distinguished differences in late ERP component, which were also reported in some of our previous studies using central presentation in eliciting a smaller LPC for both scientific metaphors and conventional metaphors as compared to literal pairs (Tang et al., 2017a). A possible reason is that late negativity overlapped in the window of LPC indicating the scientific inference for knowledge understanding.

In the LPC time window, the data revealed different hemispheric activation between scientific metaphors and conventional metaphors. Compared with literal pairs, conventional metaphors elicited significantly higher LPCs during the rvf-LH presentation, while the scientific metaphors elicited significantly lower LPCs during the lvf-RH presentation, which indicates the right hemisphere advantage in the late stage of metaphoric processing. According to the structure-mapping theory (Wolff and Gentner, 2011), the processing of metaphors involves a later directional mapping from base to target. *The career of metaphor* (Bowdle and Gentner, 2005; Wolff and Gentner, 2011) added that the bases of highly conventional metaphors already possess a salient conventional metaphoric meaning, whereas the metaphoric abstraction must be derived anew for a novel figurative. Hence a novel metaphor is understood with the analogy, while a conventional metaphor is understood with a category statement (Bowdle and Gentner, 1999, 2005), which is computationally less demanding and even easier than "automatic" (Holyoak and Stamenkovic, 2018). Combined with the fine-coarse coding hypothesis (Beeman, 2005), the mapping of conventional metaphors is simpler than the processing of literal expressions and therefore may mainly activate the left hemisphere. In contrast, the mapping of scientific metaphors may involve a comparison process to establish a new integration which is more computationally costly than literal expressions and therefore may mainly activate the right hemisphere. More importantly, scientific metaphors elicited markedly lower LPCs than conventional metaphors during the rvf-LH presentation, showing greater cognitive taxes for the left hemisphere to process analogy mapping of scientific metaphors.

In summary, the hemispheric processing of scientific metaphors might differ from that of conventional metaphors and literal expressions in the following three ways. First, the information retrieval for scientific terms involved in scientific metaphors might increase the cognitive loads of the left hemisphere. Secondly, during the processing of semantic integration and structure alignment, the distance between scientific target domains and daily source domains of a scientific metaphor might be longer than that between the two daily domains of a conventional metaphor, which might result in a higher calculative tax for the right hemisphere. Thirdly, during the later-stage processing of directional mapping, the mapping of scientific metaphors may involve a comparison process to establish a new integration which is assumed to be computationally costly for the right hemisphere. In general, the processing of scientific metaphors might activate both hemispheres and the right hemisphere plays an important role in semantic integration, structure alignment and directional mapping.

CONCLUSION

The current DVF study introduced scientific metaphors as novel metaphors presenting orientation mapping from the specific and familiar domains to the abstract and unfamiliar domains to examine hemispheric asymmetry. In addition, a scientific metaphor has a more complex context which increases the cognitive load, especially in the late stage. The results suggest that the complexity of mapping impacts the lateralization of metaphor processing. Although both hemispheres are involved in scientific metaphor comprehension, the right hemisphere takes a special role in integrating domains across a long distance and making inferences for scientific knowledge.

However, the DVF paradigm used in the study has some limitations. When a stimulus is presented on one visual field, the opposite side of the hemisphere to the visual field is activated while the same side of the hemisphere is inhibited, which might result in some counteraction of effects. This might also explain why similar studies draw conclusions that both hemifields are involved in non-literal language processing. Therefore, the extent to which the hemispheres are inhibited by hemifield presentation should be taken into data analysis is worthy of further verification. Besides, it is difficult to ensure that the stimuli presented in different visual fields would be processed by the opposite hemispheres totally due to individual physical and attentional differences. Some participants might sometimes roll their eyes unconsciously. Also, due to limited time, our study did not modulate SOA as some previous DVF experiments did. Future studies are suggested to apply different SOAs to further study the time course of non-figurative language processing, as well as to use more advanced techniques such as eye-tracking to monitor the whole experimental process.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Shaanxi Normal University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XT, MH, SX, and SH contributed to conception and design of the study. XT, MH, and SX performed the data collection. LS, MH, and YH performed the analysis. MH and LS wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was funded by the National Social Science Foundation of China (Grant No. 17BYY092), Anhui Provincial Department of Education (Grant No. SK2020A0225), and Anhui Provincial Department of Education (Grant No. TSSK2015B29).

ACKNOWLEDGMENTS

Special thanks to Michael, a British colleague at AHPU, for helping polish some of the language in our final draft.

REFERENCES

Anaki, D., Faust, M., and Kravetz, S. (1998). Cerebral hemispheric asymmetries in processing lexical metaphors. *Neuropsychologia* 36, 353–362. doi: 10.1016/S0028-3932(97)00110-3

Aurnhammer, C., Delogu, F., Schulz, M., Brouwer, H., and Crocker, M. W. (2021). Retrieval (N400) and integration (P600) in expectation-based comprehension. *PLoS One* 16:e0257430. doi: 10.1371/journal.pone.0257430

Beeman, M. J. (1998). "Coarse semantic coding and discourse comprehension," in *Right Hemisphere Language Comprehension: Perspectives From Cognitive Neuroscience*, eds M. J. Beeman and C. Chiarello (Mahwah: Lawrence Erlbaum Associates Publishers), 255–284.

Beeman, M. J. (2005). Bilateral brain processes for comprehending natural language. *Trends Cogn. Sci.* 9, 512–518. doi: 10.1016/j.tics.2005.09.009

Bohrn, I. C., Altmann, U., and Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia* 50, 2669–2683. doi: 10.1016/j.neuropsychologia.2012.07.021

Bowdle, B. F., and Gentner, D. (1999). "Metaphor comprehension: From comparison to categorization," in *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society*, eds M. Hahn and S. C. Stoness (Hillsdale: Lawrence Erlbaum Associates), 90–95. doi: 10.4324/9781410603494-21

Bowdle, B. F., and Gentner, D. (2005). The Career of Metaphor. *Psychol. Rev.* 112, 193–216. doi: 10.1037/0033-295X.112.1.193

Coulson, S., Federmeier, K. D., Van Petten, C., and Kutas, M. (2005). Right Hemisphere Sensitivity to Word- and Sentence-Level Context: evidence From Event-Related Brain Potentials. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 129–147. doi: 10.1037/0278-7393.31.1.129

Coulson, S., and Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Res.* 1146, 128–145. doi: 10.1016/j.brainres.2007.03.008

Eviatar, Z., and Just, M. A. (2006). Brain correlates of discourse processing: an fMRI investigation of irony and conventional metaphor comprehension. *Neuropsychologia* 44, 2348–2359. doi: 10.1016/j.neuropsychologia.2006.05.007

Faust, M., and Mashal, N. (2007). The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: a divided visual field study. *Neuropsychologia* 45, 860–870. doi: 10.1016/j.neuropsychologia.2006.08.010

Forgács, B., Lukács, A., and Pléh, C. (2014). Lateralized processing of novel metaphors: disentangling figurativeness and novelty. *Neuropsychologia* 56, 101–109. doi: 10.1016/j.neuropsychologia.2014.01.003

Giora, R. (1997). Understanding figurative and literal language: the graded salience hypothesis. *Cogn. Linguist.* 8, 183–206. doi: 10.1515/cogl.1997.8.3.183

Giora, R. (2003). *On Our Mind: Salience, Context, and Figurative Language*. New York: Oxford University Press.

Holyoak, K. J., and Stamenkovic, D. (2018). Metaphor comprehension: a critical review of theories and evidence. *Psychol. Bull.* 144, 641–671. doi: 10.1037/bul0000145

Ianni, G. R., Cardillo, E. R., McQuire, M., and Chatterjee, A. (2014). Flying under the radar: figurative language impairments in focal lesion patients. *Front. Hum. Neurosci.* 8:871. doi: 10.3389/fnhum.2014.00871

Kazmerski, V. A., Blasko, D. G., and Dessalegn, B. G. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Mem. Cogn.* 31, 673–689. doi: 10.3758/BF03196107

Lai, V. T., and Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain Lang.* 127, 484–496. doi: 10.1016/j.bandl.2013.09.010

Lai, V. T., Van Dam, W., Conant, L. L., Binder, J. R., and Desai, R. H. (2015). Familiarity differentially affects right hemisphere contributions to processing metaphors and literals. *Front. Hum. Neurosci.* 9:44. doi: 10.3389/fnhum.2015.00044

Lauro, L. J. R., Tettamanti, M., Cappa, S. F., and Papagno, C. (2008). Idiom Comprehension: a Prefrontal Task? *Cereb. Cortex* 18, 162–170. doi: 10.1093/cercor/bhm042

Mashal, N., and Faust, M. (2008). Right hemisphere sensitivity to novel metaphoric relations: application of the signal detection theory. *Brain Lang.* 104, 103–112. doi: 10.1016/j.bandl.2007.02.005

Mitchell, R., Vidaki, K., and Lavidor, M. (2016). The Role of Left and Right Dorsolateral Prefrontal Cortex in Semantic Processing: a transcranial direct current stimulation study. *Neuropsychologia* 91, 480–489. doi: 10.1016/j.neuropsychologia.2016.08.019

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Omoto, S., Kuroiwa, Y., Li, M., Doi, H., Shimamura, M., Koyano, S., et al. (2001). Modulation of event-related potentials in normal human subjects by visual divided attention to spatial and color factors. *Neurosci. Lett.* 311, 198–202. doi: 10.1016/S0304-3940(01)02172-3

Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., and Kircher, T. T. (2004). Neural correlates of metaphor processing. *Cogn. Brain Res.* 20, 395–402. doi: 10.1016/j.cogbrainres.2004.03.017

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., and Abraham, A. (2012). Can clouds dance? Part 2: an ERP investigation of passive conceptual expansion. *Brain Cogn.* 80, 301–310. doi: 10.1016/j.bandc.2012.08.003

Schneider, S., Rapp, A. M., Haeußinger, F. B., Ernst, L. H., Hamm, F., Fallgatter, A. J., et al. (2014). Beyond the N400: complementary access to early neural correlates of novel metaphor comprehension using combined electrophysiological and haemodynamic measurements. *Cortex* 53, 45–59. doi: 10.1016/j.cortex.2014.01.008

Sela, T., Panzer, M. S., and Lavidor, M. (2017). Divergent and convergent hemispheric processes in idiom comprehension: the role of idioms predictability. *J. Neurolinguistics* 44, 134–146. doi: 10.1016/j.jneuroling.2017.05.002

Tang, X., Qi, S., Jia, X., Wang, B., and Ren, W. (2017a). Comprehension of scientific metaphors: complementary processes revealed by ERP. *J. Neurolinguistics* 42, 12–22. doi: 10.1016/j.jneuroling.2016.11.003

Tang, X., Qi, S., Wang, B., Jia, X., and Ren, W. (2017b). The temporal dynamics underlying the comprehension of scientific metaphors and poetic metaphors. *Brain Res.* 1655, 33–40. doi: 10.1016/j.brainres.2016.11.005

Wolff, P., and Gentner, D. (2011). Structure-mapping in metaphor comprehension. *Cogn. Sci.* 35, 1456–1488. doi: 10.1111/j.1551-6709.2011.01194.x

Yang, J. (2014). The role of the right hemisphere in metaphor comprehension: a meta-analysis of functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 35, 107–122. doi: 10.1002/hbm.22160

Yang, J., Li, P., Fang, X., Shu, H., Liu, Y., and Chen, L. (2016). Hemispheric involvement in the processing of Chinese idioms: an fMRI study. *Neuropsychologia* 87, 12–24. doi: 10.1016/j.neuropsychologia.2016.04.029

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 03 June 2022
doi: 10.3389/fpsyg.2022.890962

Gesture–Speech Integration in Typical and Atypical Adolescent Readers

Ru Yao[1], Connie Qun Guan[2], Elaine R. Smolen[3], Brian MacWhinney[4], Wanjin Meng[5]* and Laura M. Morett[6]*

[1]China National Institute of Education Sciences, Beijing, China, [2]School of Foreign Studies, Beijing Language and Culture University, Beijing, China, [3]Teachers College, Columbia University, New York, NY, United States, [4]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States, [5]Department of Moral, Psychological and Special Education, China National Institute of Education Sciences, Beijing, China, [6]Department of Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, Tuscaloosa, AL, United States

This study investigated gesture–speech integration (GSI) among adolescents who are deaf or hard of hearing (DHH) and those with typical hearing. Thirty-eight adolescents (19 with hearing loss) performed a Stroop-like task in which they watched 120 short video clips of gestures and actions twice at random. Participants were asked to press one button if the visual content of the speaker's movements was related to a written word and to press another button if it was unrelated to a written word while accuracy rates and response times were recorded. We found stronger GSI effects among DHH participants than hearing participants. The semantic congruency effect was significantly larger in DHH participants than in hearing participants, and results of our experiments indicated a significantly larger gender congruency effect in DHH participants as compared to hearing participants. Results of this study shed light on GSI among DHH individuals and suggest future avenues for research examining the impact of gesture on language processing and communication in this population.

Keywords: deaf and hard of hearing, automaticity, gesture–speech integration, spoken language comprehension, gesture

OPEN ACCESS

Edited by:
Antonio Benítez-Burraco, University of Seville, Spain

Reviewed by:
Nicole Dargue, Griffith University, Australia
Naomi Sweller, Macquarie University, Australia

****Correspondence:***
Connie Qun Guan
qunguan81@163.com
Wanjin Meng
1205747017@qq.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *07 March 2022*
Accepted: *09 May 2022*
Published: *03 June 2022*

Citation:
Yao R, Guan CQ, Smolen ER, MacWhinney B, Meng W and Morett LM (2022) Gesture–Speech Integration in Typical and Atypical Adolescent Readers. Front. Psychol. 13:890962. doi: 10.3389/fpsyg.2022.890962

INTRODUCTION

With the diagnosis of hearing loss increasing in prevalence (Yuri et al., 2008; Verrecchia and Curcio, 2016) and being researched around the world (Moscicki et al., 1985; Cruickshanks et al., 1998; Reuben et al., 1998; Colozza and Anastasio, 2009), experts are debating an important question: does the loss of hearing mean simply the absence of sensory input in the auditory modality, or does it lead to enhancement of perceptual ability in other modalities, such as vision? Compensatory plasticity holds that the lack of auditory stimulation experienced by deaf individuals is accompanied by enhancements in other senses, such as visual cognition. However, some evidence in the educational and cochlear implant literature documents deficient visual cognition in individuals who are deaf or hard of hearing (DHH; Bavelier et al., 2006). Without early identification, technology use, and auditory-based intervention, hearing loss is often accompanied by difficulty developing spoken language (Taitelbaum-Swead et al., 2006), and it may also bring about disadvantages in daily life and societal discrimination (Branson and Miller, 2005). However, several studies have also suggested that the removal of one sensory modality leads to neural reorganization of the remaining modalities (Finney et al., 2001), which is regarded as compensation.

It is well established that gestures are communicative and are integrated automatically with speech by individuals with typical hearing in order to comprehend a message (Volterra and Erting, 1994; Obermeier et al., 2011, 2012). Language researchers have theorized that gesture and speech work together to form a single integrated system of meaning during language comprehension (Kendon, 1986; McNeill, 1994; Hostetter, 2011; Dargue et al., 2019; Kandana Arachchige et al., 2021). Kelly et al. (1999) pioneered research into what they termed the gesture–speech integration (GSI) effect and argued that gestures have a powerful impact on how hearing individuals comprehend and remember pragmatic communication. This automatic GSI also appears to exist in DHH individuals. Obermeier et al. (2012) conducted two experiments to investigate automaticity and the ways in which communicative abilities and the environment influence integration of gesture and speech. They found a significant benefit of using gestures during communication among the DHH group; that is, DHH participants showed better GSI and successful disambiguation of speech. Results indicated that gestures are beneficial in countering difficult communication conditions independent of whether the difficulties are due to external (ambient noise) or internal (hearing loss) factors.

Representational Gestures and Spoken Language Comprehension

In general, representational gestures, which convey meaning *via* their form and motion, facilitate comprehension of both native and non-native languages (Rogers, 1978; Beattie and Shovelton, 1999; Kelly et al., 1999, 2010b; Church et al., 2004; Holle and Gunter, 2007; Holler and Wilkin, 2011; Obermeier et al., 2011). These gestures help to disambiguate pragmatically ambiguous speech (Kelly et al., 1999) and words with ambiguous meanings (Holle and Gunter, 2007). Moreover, semantic processing is impacted by the presence of representational gestures either congruent or incongruent in meaning with co-occurring speech (Kelly et al., 2004). Moreover, they play a crucial role in language-based communication (Kelly et al., 2010b), and people extract information about meaning from them (Beattie and Shovelton, 1999). Representational gestures are taken into consideration during conversations, and one cannot avoid integrating them with speech (Kelly et al., 2010a). Representational gestures improve speech comprehension in suboptimal situations, such as in a setting with a great deal of background noise, influencing GSI (Obermeier et al., 2011; Drijvers and Özyürek, 2017). In these situations, individuals tend to regard representational gestures as a helpful and related cue for language comprehension (Rogers, 1978). Overall, representational gestures have profound influences on the processing, communication, and comprehension of spoken language.

Gesture and Speech Comprehension in the Typically Hearing Population

As discussed above, gesture, as a non-verbal disambiguation cue, influences speech interpretation among individuals with typical hearing. In particular, gestures are a relevant and helpful cue in noisy conversational settings (Rogers, 1978). According to Obermeier et al. (2011), typically hearing individuals leverage gestures to disambiguate the meaning of ambiguous speech when noise interferes with its comprehension. In the field of embodied cognition, iconic gestures, which represent physical/spatial attributes or actions, have been widely studied because of their direct connection to tangible conceptual representations (Barsalou, 2003). McNeil et al. (2000) have investigated whether these gestures support spoken language comprehension in children. They conclude that their role in speech comprehension depends on the complexity of the spoken message, and that they facilitate speech comprehension primarily for complex spoken messages. In typically hearing individuals, speech and gesture reciprocally influence one another's semantic processing during online comprehension (Ozyurek, 2010). Gesture has also been found to influence three interrelated cognitive processes sub-serving second language (L2) word learning: communication, encoding, and recall (Allen, 1995; Tellier, 2008; Macedonia et al., 2011; Macedonia and von Kriegstein, 2012; So et al., 2012; Macedonia, 2014; Morett, 2014, 2018). Above all, the effect of gesture on speech comprehension in hearing individuals depends on both the relation of gesture to speech and the complexity of the spoken message (McNeil et al., 2000), facilitating language processing and communication among native and non-native speakers.

Several studies suggest that gestures complementing spoken language in meaning facilitate learning (Sadoski, 2018; Andra et al., 2020). Porter (2016) and Tellier (2008) examined the effects of gesture production on children's L2 vocabulary acquisition. Tellier (2008) demonstrated that production of gestures conveying the meanings of L2 words may facilitate children's L2 vocabulary learning, whereas Porter (2016) shows that such gestures may facilitate it for 5- and 6-year-old when combined with images of referents. Andra et al. (2020) compared the effect of gestures and images depicting word referents on children's L2 vocabulary learning. They conclude that gesture significantly benefits L2 vocabulary learning in comparison with learning without gesture, whereas it does not significantly benefit it in comparison with images. Further, the effects of gesture on L2 vocabulary learning last for several months, indicating that it facilitates long-term memory for L2 words.

Gestures also facilitate the understanding of abstract concepts, and different types of gestures have different effects on it (Kang et al., 2013). Moreover, gestures benefit comprehension of spoken narratives (Hough, 1990; Lyle, 2000; Schmithorst et al., 2006; Dargue and Sweller, 2020). Finally, there is growing evidence that gestures can enhance acquisition of novel L2 speech sounds (Morett and Chang, 2015; Zheng et al., 2018; Baills et al., 2019; Zhen et al., 2019; Hoetjes and Van Maastricht, 2020; Xi et al., 2020; Morett et al., 2022).

Gesture Comprehension and Use in the DHH Population

Many representational gestures approximate signs from signed languages, which may aid communication between individuals who are DHH and those with typical hearing. Kendon (1997) considered gesticulations (i.e., co-speech gestures), emblems, and signs all to be gestures, but did not consider posture shifts, self-adaptors (e.g., grooming, scratching), and object manipulations to be gestures. It is likely that DHH individuals have permanently adapted their communicative systems to incorporate as much extra-linguistic information as possible (Obermeier et al., 2012). Deaf people often fixate visually on the face to pick up microexpressions and movements of the articulators (Muir and Richardson, 2005). Thus, individuals experiencing difficulties hearing speech tend to use available visual information to improve their speech comprehension (Muir and Richardson, 2005). This makes co-speech gestures a powerful tool to support speech comprehension in daily communication.

Comprehension of co-speech gestures has been researched more extensively in DHH individuals than comprehension of other gesture types because of their supporting role in language processing (Krauss et al., 1991; McNeill, 1994; Kelly et al., 2010b). In oral-deaf individuals who tend to be born to hearing parents and learn to communicate orally and to read words on the lips of the speakers (Vendrame et al., 2010), gestures accompanying discourse facilitate retention of content information and correct inferences. However, co-speech gestures interfere with verbatim memory for discourse in these individuals. In addition, gestures produced by DHH individuals also complement sign language; for example, gestures can be used to request a turn during a sign language conversation (Emmorey, 1999).

GSI in the Hearing and DHH Populations

GSI is an automatic process supporting language comprehension. When a semantically incongruent gesture–speech combination is presented, processing of gesture is negatively affected by incongruent speech, and processing of speech is also negatively affected by incongruent gesture. That is, concurrent speech and gestures influence each other's processing (Kelly et al., 2010b). Kelly et al.'s (1999) seminal work proposed the GSI effect and argued that gestures have a powerful impact on how speech is comprehended and remembered. Kelly et al. (2010b) further explored the strength of the neural relationship between gesture and speech by examining a potential interface between language and action in the brain (i.e., GSI). Through a Stroop-like task, this study provided evidence supporting the GSI effect. When participants' attention was drawn to the semantic relationship between speech and gesture, a larger N400 effect (which indexes semantic integration, as in Kutas and Hillyard, 1980) was observed when spoken words were accompanied by semantically incongruent vs. congruent gestures. Zhao et al. (2018) explored the neurocognitive control mechanism of GSI using transcranial magnetic stimulation (TMS), showing that disrupting activity in related brain regions (inferior frontal gyrus or posterior middle temporal gyrus) selectively impairs GSI.

Language comprehension is also influenced by gesture in GSI tasks in DHH individuals (Obermeier et al., 2011, 2012). Obermeier et al.' (2012) seminal study examined the GSI effect in DHH individuals to determine whether gestures influence their comprehension to a greater extent than that of hearing individuals. They found that spoken language comprehension in DHH individuals is heavily influenced by gesture, like typically hearing subjects in the noisy condition. It seems that individuals with normal hearing adapt their gesture production and comprehension based on the quality of the auditory speech signal, whereas DHH individuals have permanently adapted their communication to incorporate as much extra-linguistic information as possible, leading them to incorporate gesture with spoken language with greater automaticity.

Cross-Modal Plasticity and Multimodal Integration in the DHH Population

DHH and hearing individuals differ in visual cognitive ability, spatial distribution of attention to the peripheral field, and multimodal reorganization. Enhancements in visual cognition have been noted in DHH individuals in comparison with hearing individuals when confounding variables are controlled. These changes are limited to aspects of vision that are attentionally demanding and benefit from auditory–visual convergence (Bavelier et al., 2006). Moreover, deafness appears to shift the spatial distribution of attention such that attention to the peripheral, but not the central, visual field is heightened (Bavelier et al., 2000, 2001, 2006; Obermeier et al., 2012). When asked to detect the direction of motion of a peripherally located stimulus, deaf individuals do so more quickly and accurately than hearing individuals (Neville and Lawson, 1987). Furthermore, effective connectivity between middle temporal (MT)/middle superior temporal (MST) and posterior parietal cortex is stronger in deaf than hearing individuals during peripheral but not central attention. Greater sensitivity to peripheral motion enables deaf individuals to process large, swift hand movements like signs and gestures efficiently even when focusing attention on the interlocutor's face (Muir and Richardson, 2005). Thus, hearing loss may, to some extent, facilitate visual skills in DHH individuals compared with hearing individuals. Visual stimuli activate the auditory cortex in deaf individuals who sign, suggesting that the removal of one sensory modality in humans leads to neural reorganization of the remaining modalities, at least for those who use signed language (Finney et al., 2001). A common feature of functionally reorganized brain areas in DHH individuals is their role in multimodal processing, reinforcing recent views on the importance of multimodal integration at all stages of cognitive processing (Ghazanfar and Schroeder, 2006). Overall, these results suggest that cross-modal plasticity may serve as a core compensatory mechanism *via* enhanced modulation of spatial attention in the visual modality (Eimer et al., 2002).

Present Study and Hypotheses

Previous research has demonstrated that language is linked to action *via* gesture (Willems, 2007; Willems and Hagoort, 2007;

Hostetter and Alibali, 2010). The present study investigates the strength of this relationship and extends it to adolescents by focusing on a potential interface between two systems: representational gestures and speech.

First, considering evidence that gesture is similarly semantically related to speech and text (Hughes-Berheim et al., 2020), we compared the GSI effect across the visual and auditory modalities to examine how gesture is integrated with language in both modalities.

Second, we investigated the automaticity of GSI by using a modified Stroop task in which participants were asked to identify the speaker's gender (a superficial task) to decrease attention on semantic congruency or incongruency between the prime and target (a goal task), following Kelly et al. (2010a).

Third, the GSI effect was compared in DHH and hearing adolescents to explore whether DHH individuals experience greater automaticity in integrating gestures and speech than the typical hearing group.

We hence hypothesized that (1) the GSI effect would not show a significant difference between auditory and visual modalities for either the DHH or the hearing group due to the DHH group's use of assistive hearing technology (i.e., hearing aids or cochlear implants); (2) both DHH and hearing adolescents would respond slower to gestures and speech when they were incongruent compared to when they were congruent; and (3) the GSI effect would be stronger in DHH than hearing adolescents indicating DHH individuals experience greater automaticity in integrating gestures and speech than typical hearing individuals. Because the experimental conditions for both groups were identical, consisting of acoustically and visually clear recordings, this should eliminate the potentially confounding effects of background noise and other distractions, allowing any differences between the DHH and hearing groups to be attributed to differences in their processing of gestures accompanying spoken and written language.

MATERIALS AND METHODS

Participants

Thirty-eight native speakers of Chinese provided written informed consent to participate in the current study. All participants were all right-handed, had normal or corrected-to-normal vision, had no known neurological deficits, and had not taken part in a similar experiment using the same stimuli. Nineteen participants (10 females; $M_{age} = 13$ years, age range: 11–15 years) were DHH. We recruited adolescents as the target population because relatively little research has been conducted on GSI in adolescents (Dargue et al., 2019). The hearing and DHH groups were age matched. Seventy percent of the DHH participants had mild-to-moderate hearing loss (unaided pure tone average range: 70–115 dB); the other 30% had severe to profound hearing loss (unaided pure tone average range <70 dB; Baille et al., 1996). Most DHH participants used sign language in daily communication. Only two participants reported that they used spoken language in their interactions with the general population. DHH participants were recruited from educational programs where spoken Chinese and Chinese Sign Language were used simultaneously by teachers and students throughout the day. All DHH participants communicated in spoken Chinese during the experiment, and all used assistive hearing technology (i.e., hearing aids or cochlear implants) to access speech sounds during the experiment. Eighty-five percent of the DHH participants had hearing parents. Nineteen participants (nine females; $Mean_{age} = 13.45$ years, age range: 13–18 years) had typical hearing. Hearing levels of participants in the control group were tested using an ISO-audiogram with the frequency bands 500, 1,000, 2,000, and 4,000 Hz (Moore, 2014). They had a mean hearing level of 13 dB (range: 6–19 dB), which is well within the 25 dB range typically defined as the boundary for normal hearing (Bies and Hansen, 2003). All participants were paid $10 USD for their participation.

Materials

Participants viewed a randomized sequence of 120 short video clips twice (once in congruent trials, and once in incongruent trials). A semantically related or unrelated prime was shown on the screen following a video clip (interleaved across trials), for a total of 240 presentations. Stimuli were divided across two groups, such that half of the trials were "related" (i.e., congruent) and the other half were "unrelated" (i.e., incongruent). For example, referring to **Figure 1**, suppose that the primes were the words *dial* and *cut*. Panels A and B are related (congruent), whereas Panels C and D are unrelated (incongruent), creating a completely balanced design. Participants were instructed to press one button if the gesture related to the prime and another button if it did not relate to the prime. The experimental procedure lasted approximately 20 min.

Stimuli/Materials

Digital videos were created with a Sony DV 100 digital camcorder and edited with Final Cut Pro software. Videos showed a male or female actor situated in natural contexts (e.g., kitchen, living room, entryway) describing everyday activities (e.g., drinking water, watering, tying shoes). We used different backgrounds to increase the ecological validity of the results. The actors faced the camera in all videos, but their faces were digitally covered so that mouth movements were not visible. Subjects were told that this was to hide the actors' identities. Actors spoke at a normal pace with no artificial pauses between words.

Forty-four digitized videos of iconic gestures (e.g., break, twist) were selected for use in the current study based on previous studies (Kelly et al., 2010a; Dick et al., 2014). Each gesture was produced by either a male or a female while simultaneously uttering the corresponding verb in Chinese (**Figure 1**). In a follow-up session, the two speakers were recorded producing words only as speech. Video and audio materials were then combined to create the experimental manipulations of gender and semantic congruency. The presentation software application was used to present stimuli to participants, and response buttons were counterbalanced. Half of materials required responses of "left" for male and "right" for female, and the other half were vice versa.

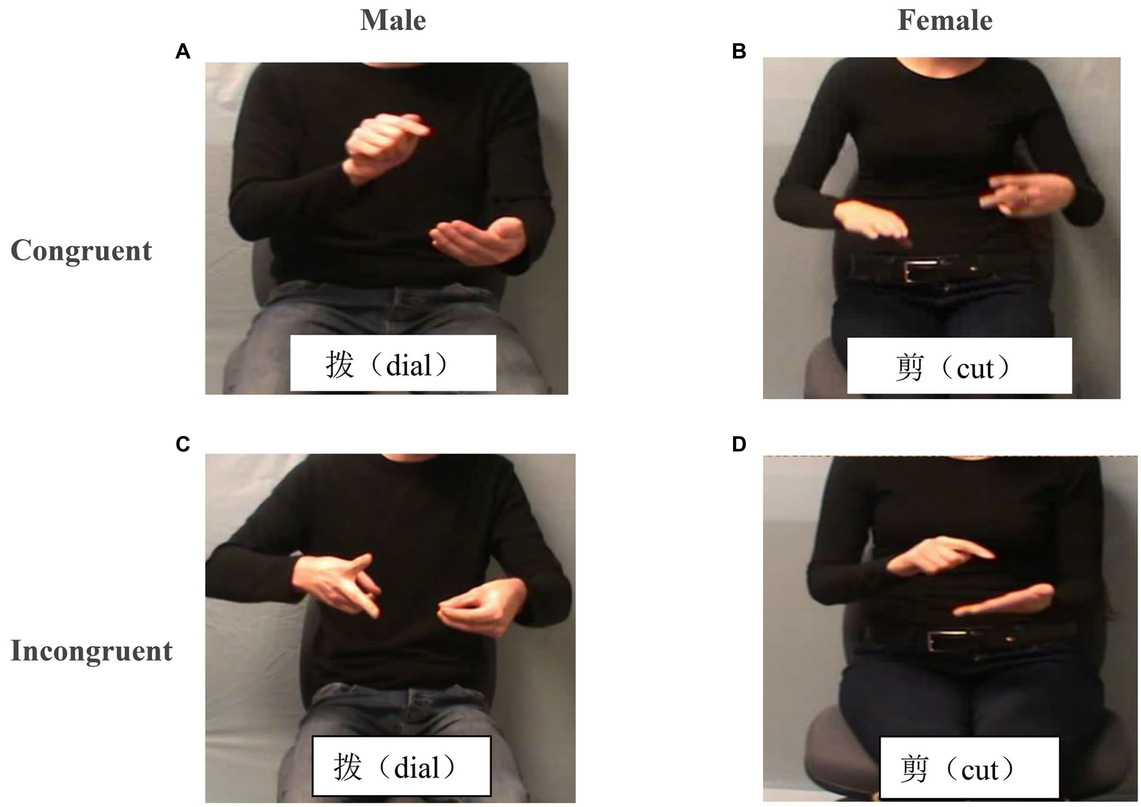

FIGURE 1 | Semantically congruent and incongruent Stroop-like gesture–speech integration (GSI) task paradigm. Illustrations of the stimuli. In the gesture-congruent condition (picture **A**, **B**), the gesture and the speech were semantically related (e.g., gesturing dial and saying "*dial*" or gesturing cut and saying "*cut*"), and in the gesture-incongruent condition (picture **C**, **D**), they were semantically unrelated (e.g., gesturing cut and saying "*dial*," and vice versa). There was a gender manipulation that varied the relationship between the gesturer and the speaker. In the gender-same condition, the gesturer and the speaker were the same person (a man or a woman), but in the gender-different condition, the gesturer and the speaker were two different people (e.g., a man gesturing but a woman speaking, and vice versa).

Prior to presentation of videos, a written Chinese word was displayed on the screen, serving as the prime in the task (see below). The word was displayed for 500 ms, followed by a blank screen for 500 ms prior to stimulus onset. Each word displayed was an action verb used in one of the experimental conditions, and it was either related or unrelated to the auditory and/or the visual information presented in the video. The variable intertrial interval between each prime-target pair ranged from 1.5 to 2.5 s at random following Kelly et al. (2010a) and Guan and Fraundorf (2020).

Prior to the experiment, participants' classroom teachers pretaught all the vocabulary used in the task to ensure familiarity with action verbs and their meanings. This familiarity training was designed to ensure that any differences in response times were due to semantic congruency rather than receptive vocabulary knowledge (Guan et al., 2019).

Validation of Experimental Materials and Procedures

Semantic Congruency Norming

To verify the semantic congruency of gesture – speech combinations, a separate set of hearing participants ($n = 30$) rated the relationship between gesture and speech in each video on a five-point Likert scale (1 = no relation, 5 = very strong relation). The mean rating for congruent videos was 4.82 ($SD = 0.41$), whereas the mean rating for the incongruent videos was 1.21 ($SD = 0.28$), differing significantly between groups ($t = 5.11$, $p < 0.001$).

Validation of Paradigm and Stimulus Set

The RT paradigm by which participants indicated whether the stimuli were congruent or incongruent was validated in hearing participants with similar language backgrounds and abilities. The findings of Kelly et al. (2010a) were replicated, validating the stimulus set and procedures in both the DHH and the typical hearing groups.

Procedures

We used a Stroop-like paradigm (Kelly et al., 2010a) to test GSI. The classic Stroop technique presents color words in different colored fonts, and the Stroop effect arises when the meaning of the written word influences how quickly and accurately the color of the font can be named (Stroop, 1935). We used a modified version of the classic Stroop procedure in which we asked participants to judge the gender of the voice of the speaker in the video as a superficial task to avoid explicitly drawing attention to gesture and speech, which may unintentionally encourage conscious and strategic processing of the two modalities. The gender congruency task was also examined to explore the automaticity of GSI effect. Because standard Chinese was used in all stimuli, dialectical differences should not have influenced gesture processing.

In the modified Stroop task, participants responded as quickly and as accurately as possible by pressing a button to indicate whether the voice in the video was a male or a female. Each video started with the onset of a gesture stroke, with speech onset occurring 200 ms later. Practice trials were provided to ensure that all participants reached 100% accuracy on gender judgments. Accuracy was at ceiling in this task; therefore, it was not analyzed. The 7.1% of trials with errors in gender judgment were excluded from RT analyses. RTs were calculated relative to spoken word onset. Outliers were defined as RTs 2.5 or more SDs outside of each individual participant's mean RT. Overall, this resulted in 8.2% of trials being excluded as outliers, within the 5%–10% region recommended by Ratcliff (1993).

Design and Analyses

A 2 (modality, auditory vs. visual) x 2 (semantic congruency, congruent vs. incongruent) x 2 (gender congruency, congruent vs. incongruent) x 2 (group, DHH vs. hearing/control) repeated-measures ANOVA was conducted with modality, semantic congruency, and gender congruency as within-participant factors, group as a between-participant factor, and RT as a dependent variable.

Above all, we assessed the effect of modality on GSI across groups. Then, we further assessed the overall GSI effect by collapsing across groups. To manipulate gender congruency, the genders of the voice and the actor in gesture videos were counterbalanced to either match or mismatch. This is a key characteristic of the Stroop-like task introduced by Kelly et al. (2010a). Therefore, the first step of analyses was conducted to reveal the automaticity of GSI by assessing the gender congruency effect. To manipulate semantic congruency, a gesture was paired with a semantically incongruent speech token (e.g., gesturing ironing while saying "whisk"). Importantly, the reverse combination was also presented (e.g., gesturing whisking while saying "iron"), ensuring that item-specific effects were counterbalanced in modality across the stimulus set. The goal of the experiment was to test sensitivity to semantic congruency, but the superficial task requirement was to indicate whether the voice of the speaker was male or female by pressing the corresponding button.

When interactions between congruency and group reached significance, we examined GSI effects separately in the DHH and typical hearing groups, and we conducted planned orthogonal *t*-tests (two-tailed) to determine the effect sizes for audio and visual target stimuli.

RESULTS

We excluded all incorrect and skipped trials from the data, as well as outliers (±2.5 SD). First, a 2×2×2×2 repeated-measures ANOVA performed on RT data, with semantic congruency (congruent, incongruent), gender congruency (congruent, incongruent), modality (auditory, visual) as within-participant factor and group (DHH, control) as between-participant factor, was conducted to examine the main effect of modality and modality x semantic congruency x gender congruency interaction effect. In order to answer the three research questions, we would first examine the main effect of modality and the interaction effects involved with the modality to test the first null hypothesis of modality. Then we tested the second hypothesis of the overall GSI effect by examining the interaction effects of semantic congruency by gender congruency in both groups. Finally, to test the group difference hypothesis, we conducted the simple main effects between groups (typical hearing vs. DHH).

A 2 (modality, auditory and visual) x 2 (semantic congruency, congruent and incongruent) x 2 (gender congruency, congruent and incongruent) x 2 (group, DHH and hearing/control) repeated-measures ANOVA was performed on RT data using semantic congruency, gender congruency, and modality as within-participant factors and group as a between-participant factor. There was no significant main effect for modality [F (1,38) = 2.151, $p = 0.131$, $\eta^2 = 0.03$] and no significant modality x gender congruency x semantic congruency interaction effect [F (1,38) = 2.177, $p = 0.108$, $\eta^2 = 0.03$], indicating that the effect of gender congruency does not vary by modality. **Table 1** shows the RTs across the gender and semantic conditions in two modalities across groups. There were no differences in effect size in the overall GSI effect between audio vs. visual target stimuli. See the marginal means between the two modalities in **Table 1**.

To test the overall GSI effect, the 2 (semantic congruency, congruent and incongruent) x 2 (gender congruency, congruent and incongruent) x 2 (group, DHH and hearing/control) repeated-measures ANOVA performed on RT data revealed the main effect of gender congruency reached significance, F (1,38) = 37.271, $p < 0.001$, $\eta^2 = 0.08$. RTs were longer when the gender of the spoken voice and the speaker in the video were incongruent (M = 656 ms, SE = 15 ms) compared to when they were congruent (M = 638 ms, SE = 16 ms); and the simple effect of semantic congruency is significant, $F(1,39) = 5.327$, $p = 0.026$, $\eta^2 = 0.11$, indicating longer RTs when gesture and speech were

TABLE 1 | RTs across the gender and semantic conditions in two modalities across groups.

	Audio		Visual		Marginal mean
	SC	SI	SC	SI	
Gender same	621 (98)	660 (86)	614 (92)	647 (101)	636
Gender different	652 (102)	678 (98)	636 (101)	650 (102)	656
Marginal mean	635	659	624	648	----

sc, semantic congruent and si, semantic incongruent. SDs were presented in the parenthesis.

incongruent ($M = 650$ ms, $SE = 15$ ms) compared to when they were congruent ($M = 630$ ms, $SE = 15$ ms).

Simple main effects revealed that, for the between-participant group factor (typical hearing vs. DHH), there was a significant effect of gender congruency (F (1,38) = 38.12, $p < 0.001$, $\eta^2 = 0.17$), and a significant effect of semantic congruency, F (1,38) = 46.17, $p < 0.001$, $\eta^2 = 0.22$. Specifically, in the DHH group, RTs were longer when the gender of the spoken voice and the individual in the video were incongruent ($M = 658$ ms, $SE = 12$ ms) compared to when they were congruent ($M = 625$ ms, $SE = 12$ ms) with a marginal means around 33 ms, whereas this was not the case in the control group ($M = 651$ ms, $SE = 12$ ms for incongruent trials, and $M = 634$ ms, $SE = 12$ ms for congruent, with a difference around 17 ms). Moreover, in the DHH group, RTs were longer when gestures and speech were incongruent ($M = 639$ ms, $SE = 12$ ms) compared to when they were congruent ($M = 585$ ms, $SE = 12$ ms) with a difference around 64 ms, whereas this was not the case in the typical hearing control group ($M = 630$ ms, $SE = 12$ ms for incongruent trials, and $M = 592$ ms, $SE = 12$ ms for congruent, with a difference around 38 ms). **Figure 2** shows the marginal means of these simple effects on RTs for semantic congruency and gender congruency by group. All significance levels were smaller than 0.05.

DISCUSSION

In the current study, we examined whether GSI differed between hearing and DHH adolescents to investigate how it relates to reading. Using a Stroop-like lexical decision task with visually presented Chinese characters (e.g., 剪 "cut" or 拨 "dial") with speech consistent or inconsistent with the meanings of gestures, participants were asked to decide whether speakers were male or female. Responses in this task indicated high levels of automaticity among all participants, showing nearly 100% accuracy, but varied in RT between conditions. Three major findings were obtained. First, we found GSI in both DHH and hearing participants did not differ in the visual and auditory modalities, which are both important for language processing. Second, we found that automaticity in GSI among DHH participants differed by semantic and gender congruency in both the audio and visual conditions. Third, a comparison between DHH and hearing adolescents suggested differences in the magnitude of semantic and gender congruency effects between the two groups. There were significantly larger effects of semantic congruency and gender congruency in DHH participants compared to hearing participants, with a greater difference for semantic than for gender congruency. In other words, incongruency of the visual and auditory modalities more negatively influenced GSI among hearing participants in comparison with DHH participants. Thus, modality may not influence language processing in DHH individuals to the extent that it does in hearing individuals. To conclude, GSI in the DHH individuals is not restricted by modality, at least in our experimental task.

Cross-Modal Plasticity of GSI Among DHH Individuals

Our study is one of few empirical studies establishing the GSI effect in DHH and hearing individuals. We not only revealed automaticity of GSI among DHH individuals, but also found that it was unaffected by modality. Previous research suggests that when the brain is deprived of input from one sensory modality, this loss is often compensated in one or more intact sensory systems. In this way, DHH individuals may compensate for decreased auditory input through visual processing, and this compensation may contribute to GSI automaticity. For instance, for DHH individuals

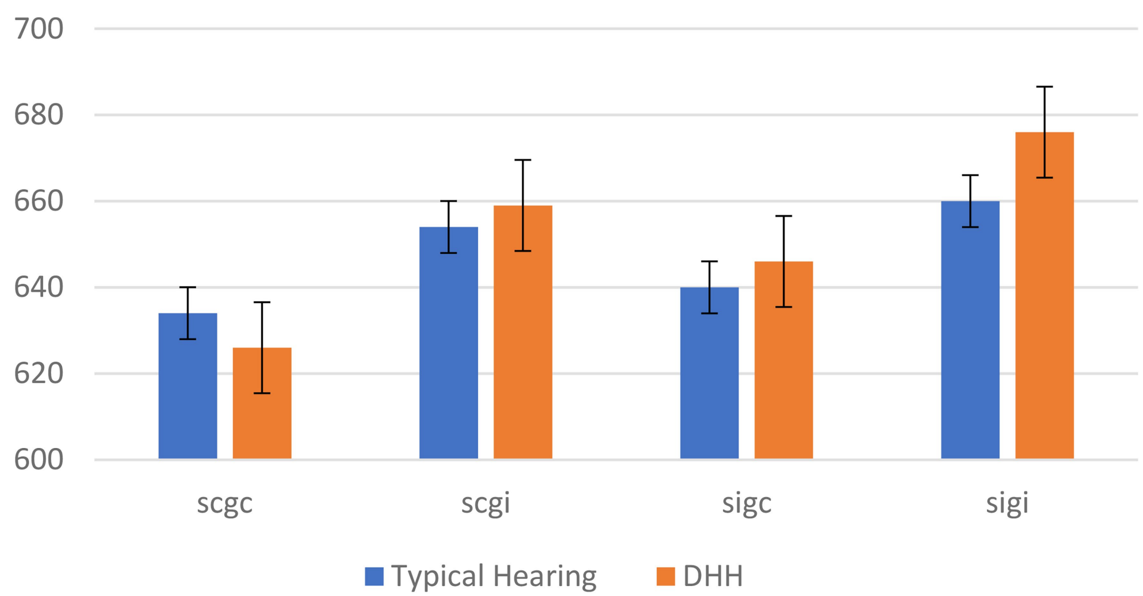

FIGURE 2 | RTs for semantic congruency and gender congruency between groups. scgc, semantic congruent and gender congruent; scgi, semantic congruent and gender incongruent; sigc, semantic incongruent and gender congruent; and sigi, semantic incongruent and gender incongruent.

who do not experience acoustic input, it has been shown that cross-modal reorganization of auditory cortex might provide compensatory visual function (Lomber et al., 2010). In other words, individuals with hearing loss are more likely to take advantage of available visual information and to regard this information as a default strategy to enhance their speech understanding.

DHH individuals recruit a special brain region, the motion processing area in middle temporal lobe, during peripheral attention after deafness (Bavelier et al., 2001). Meanwhile, DHH signers show increased activation of the posterior parietal cortex, supporting the view that parietal functions are modified after early auditory deprivation. Bavelier et al. (2001) studied the impact of early auditory deprivation on the organization of neural systems for visual motion processing and suggested that the polymodal area was modified after early sensory deprivation. This polymodal area refers to the dorsal "where" visual pathway projecting to the parietal cortex and is specialized for the perception of motion and for the localization of objects (Bavelier et al., 2001).

Our results provide evidence that visual attention in DHH individuals is comparable to that of hearing individuals. This finding is supported by Bavelier et al. (2000), who compared congenitally deaf and hearing individuals' monitoring of moving stimuli occurring in the center of the visual field and found that deaf individuals devote more attention to peripheral visual space. Bavelier et al. (2006) explored whether DHH individuals had better visual skills, observing enhanced visual cognition in this population. Importantly, auditory deprivation was associated with enhanced peripheral, compared with central, visual attention. Furthermore, Finney et al. (2001) illustrated that visual stimuli were processed in the auditory cortex in deaf individuals, providing evidence that impoverished auditory input brings about neural reorganization of visual processing. Similarly, Proksch and Bavelier (2002) demonstrated that deaf individuals attended more to the visual periphery and less to the center compared to hearing individuals.

Automaticity of GSI in DHH Individuals

Participants were slower to judge the gender of the speaker when gesture and speech were semantically incongruent, even though the semantic relationship between gesture and speech was not relevant to the task. The RT cost incurred by semantically incongruent gesture–speech pairs suggests that the representational content of gesture is automatically integrated with the representational content of speech. However, incongruent gesture–speech combinations elicited larger reaction time costs in DHH participants as compared to hearing participants, suggesting greater automaticity of GSI in DHH participants.

Some might predict that DHH individuals cannot distinguish differences in semantic congruence in gesture–word pairs when presented with gestures and semantically matching or mismatching Chinese characters. However, the results of the current study demonstrate that an automatic GSI effect exists among DHH adolescents. Even though the magnitude of the GSI effect was not as large in hearing participants as in DHH, the difference in the GSI effect between the two groups was marginal. The GSI effect in DHH individuals indicates that they integrate the semantics of gestural cues with speech.

This finding is consistent with previous research examining cognitive strategies that DHH individuals use to improve speech comprehension. For instance, Obermeier et al. (2012) found that DHH participants compensated for hearing loss by incorporating as many gestural cues as possible to improve speech comprehension. Our experimental stimuli were inspired by the stimuli of Holle and Gunter (2007), gesture fragments from Grosjean (1996), and multi-speaker babble speech created by overplaying speech streams used by Kelly et al. (2010a). Based on our behavioral data and ERP evidence from other researchers, some important conclusions can be drawn by comparing the GSI competence of DHH participants with age-matched hearing controls.

First and foremost, it appears that DHH participants tended to take gestures into account to a greater extent than hearing participants. Secondly, gestures were immediately taken into consideration by DHH participants, as their average response time in the congruent condition was quicker than in the incongruent condition. Thirdly, the accuracy rates of DHH participants in completing the experimental task were identical to hearing participants, suggesting that DHH participants may have embraced visual cues to compensate for impoverished hearing (Muir and Richardson, 2005).

For people with sensory disabilities, such as DHH, the processing of information using unaffected senses may be strengthened *via* compensation. Thus, speculatively speaking, GSI may be more efficient in DHH participants than hearing participants. In the development of hearing individuals, *skilled suppression* may be employed, resulting in slower and possibly more accurate information processing (Gernsbacher and Faust, 1995). This mechanism may have resulted in implicit GSI among hearing participants in our experiment. This conclusion requires further research, however, as we discuss in the following section.

Limitations and Future Research

The present study has several limitations. One such limitation is the sample size. Our study included 38 participants in total, of which 19 were DHH. Further, our participants were restricted in age. We only recruited adolescents (aged 11–15), so the extent to which our results generalize to other age groups is unclear. Although our research materials were ecologically valid, the differences in natural background may have distracted participants.

There are many avenues for future research in this area. First, the cognitive processes underpinning gesture, language, and writing for DHH individuals are still underinvestigated. Guan et al. (2019), for example, researched the relations between sign language and Chinese character handwriting patterns among deaf children and revealed that their semantic priming in sign language was well-integrated with their semantic priming *via* hand writing of Chinese characters. Lexical items consisted of a finite set of hand shapes, spatial locations, and movements in sign language. Similarly, Chinese characters were made up of a finite set of radicals and forms, and these units were also spatially and visually connected to one another in writing. However, research studies on the relation between gesture and Chinese character patterns among DHH children are few and far between. Second, most studies with DHH populations have been conducted in a monolingual rather than a bilingual environment. Comparisons between DHH populations using two different written languages are rare. For instance, the question of whether differences between Chinese characters and English letters influence language processing in DHH readers remains unanswered. Third, the mechanisms used by DHH individuals during multimodal language processing remain theoretical and require more empirical study. Such research would have important implications for educators as they work to develop language (signed, spoken, or both) and literacy skills for DHH children and adolescents.

Conclusion

A Stroop-like GSI task was used to compare the automatic GSI effect among adolescents who are DHH and those with typical hearing. Results suggested found that automaticity in GSI among DHH participants differed by semantic congruency in both the audio and visual conditions. There were significantly larger effects of semantic congruency and gender congruency in DHH participants compared to hearing participants. Meanwhile, the incongruency of the visual and auditory modalities more negatively influenced GSI among hearing participants in comparison with DHH participants. To conclude, GSI in the DHH individuals is not restricted by modality, at least in our experimental task.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Beijing Language and Culture University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

RY and CQG designed and conducted the study. CQG and WM analyzed the data. CQG and ERS wrote the paper. BM and LM commented and revised the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the China National Social Science Funds (#BBA180075) awarded to RY as a Principal-Investigator (PI), and CQG as a Co-PI.

REFERENCES

Allen, L. Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *Mod. Lang. J.* 79, 521–529. doi: 10.1111/j.1540-4781.1995.tb05454.x

Andra, A., Mathias, B., Schwager, A., Macedonia, M., and von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children. *Educ. Psychol. Rev.* 32, 815–850. doi: 10.1007/s10648-020-09527-z

Baille, M. F., Arnaud, C., Cans, C., Grandjean, H., du Mazaubrun, C., and Rumeau-Rouquette, C. (1996). Prevalence, aetiology, and care of severe and profound hearing loss. *Arch. Dis. Child.* 75, 129–132. doi: 10.1136/adc.75.2.129

Baills, F., Suárez-González, N., González-Fuente, S., and Prieto, P. (2019). Observing and producing pitch gestures facilitates the learning of mandarin Chinese tones and words. *Stud. Second. Lang. Acquis.* 41, 33–58. doi: 10.1017/s0272263118000074

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Lang. Cogn. Process.* 18, 513–562. doi: 10.1080/01690960344000026

Bavelier, D., Brozinsky, C., Tomann, A., Mitchell, T., Neville, H., and Liu, G. (2001). Impact of early deafness and early exposure to sign language on the cerebral organization for motion processing. *J. Neurosci.* 21, 8931–8942. doi: 10.1523/jneurosci.21-22-08931.2001

Bavelier, D., Dye, M. W. G., and Hauser, P. C. (2006). Do deaf individuals see better? *Trends Cogn. Sci.* 10, 512–518. doi: 10.1016/j.tics.2006.09.006

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., et al. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *J. Neurosci.* 20:RC93. doi: 10.1523/JNEUROSCI.20-17-j0001.2000

Beattie, G., and Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica* 123, 1–30. doi: 10.1515/semi.1999.123.1-2.1

Bies, D. A., and Hansen, C. H. (2003). *Engineering Noise Control: Theory and Practice. 3rd Edn.* London, United Kingdom: Chapman and Hall.

Branson, J., and Miller, D. (2005). *Damned for Their Difference: The Cultural Construction of Deaf People as Disabled*, vol. *7*. Washingtong DC: Gallaudet University Press, 129–132.

Church, R. B., Ayman-Nolley, S., and Mahootian, S. (2004). The role of gesture in bilingual education: does gesture enhance learning? *Int. J. Biling. Educ. Biling.* 7, 303–319. doi: 10.1080/13670050408667815

Colozza, P., and Anastasio, A. R. T. (2009). Screening, diagnosing and treating deafness: the knowledge and conduct of doctors serving in neonatology and/or pediatrics in a tertiary teaching hospital. *São Paulo Med. J.* 127, 61–65. doi: 10.1590/s1516-31802009000200002

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E. K., Klein, R., Mares-Perlman, J. A., et al. (1998). Prevalence of hearing loss in older adults in beaver dam, Wisconsin: the epidemiology of hearing loss study. *Am. J. Epidemiol.* 148, 879–886. doi: 10.1093/oxfordjournals.aje.a009713

Dargue, N., and Sweller, N. (2020). Learning stories through gesture: gesture's effects on child and adult narrative comprehension. *Educ. Psychol. Rev.* 32, 249–276. doi: 10.1007/s10648-019-09505-0

Dargue, N., Sweller, N., and Jones, M. P. (2019). When our hands help us understand: a meta-analysis into the effects of gesture on comprehension. *Psychol. Bull.* 145, 765–784. doi: 10.1037/bul0000202

Dick, A. S., Mok, E. H., Raja, B. A., Goldin-Meadow, S., and Small, S. L. (2014). Frontal and temporal contributions to understanding the iconic cospeech gestures that accompany speech. *Hum. Brain Mapp.* 35, 900–917. doi: 10.1002/hbm.22222

Drijvers, L., and Özyürek, A. (2017). Visual context enhanced: the joint contribution of iconic gestures and visible speech to degraded speech comprehension. *J. Speech Lang. Hear. Res.* 60, 212–222. doi: 10.1044/2016_jslhr-h-16-0101

Eimer, M., Velzen, J. V., and Driver, J. (2002). Cross-modal interactions between audition, touch, and vision in endogenous spatial attention: ERP evidence on preparatory states and sensory modulations. *J. Cogn. Neurosci.* 14, 254–271. doi: 10.1162/089892902317236885

Emmorey, K. (1999). “Do signers gesture?” in *Gesture, Speech, and Sign*. eds. L. Messing and R. Campbell (Oxford: Oxford University Press), 133–159.

Finney, E. M., Fine, I., and Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat. Neurosci.* 4, 1171–1173. doi: 10.1038/nn763

Gernsbacher, M. A., and Faust, M. (1995). “Skilled suppression,” in *Interference and Inhibition in Cognition*. ed. F. N. Denster (London, United Kingdom: Academic Press), 295–327.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285. doi: 10.1016/j.tics.2006.04.008

Grosjean, F. (1996). Gating. *Lang. Cogn. Process.* 11, 597–604. doi: 10.1080/016909696386999

Guan, C. Q., and Fraundorf, S. H. (2020). Cross-linguistic word recognition development among Chinese children: a multilevel linear mixed-effects modeling approach. *Front. Psychol.* 11:544. doi: 10.3389/fpsyg.2020.00544

Guan, C. Q., Zhao, J., Kwok, R. K. W., and Wang, Y. (2019). How does morphosyntactic skill contribute to different genres of Chinese writing from grades 3 to 6? *J. Res. Read.* 42, 239–267. doi: 10.1111/1467-9817.12239

Hoetjes, M., and Van Maastricht, L. (2020). Using gesture to facilitate L2 phoneme acquisition: the importance of gesture and phoneme complexity. *Front. Psychol.* 11:575032. doi: 10.3389/fpsyg.2020.575032

Holle, H., and Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *J. Cogn. Neurosci.* 19, 1175–1192. doi: 10.1162/jocn.2007.19.7.1175

Holler, J., and Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *J. Nonverbal Behav.* 35, 133–153. doi: 10.1007/s10919-011-0105-6

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychol. Bull.* 137, 297–315. doi: 10.1037/a0022128

Hostetter, A. B., and Alibali, M. Q. (2010). Language, gesture, action! A test of the gesture as simulated action framework. *J. Mem. Lang.* 63, 245–257. doi: 10.1016/j.jml.2010.04.003

Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: theme organization. *Brain Lang.* 38, 253–277. doi: 10.1016/0093-934x(90)90114-v

Hughes-Berheim, S. S., Morett, L. M., and Bulger, R. (2020). Semantic relationships between representational gestures and their lexical affiliates are evaluated similarly for speech and text. *Front. Psychol.* 11:2808. doi: 10.3389/fpsyg.2020.575991

Kandana Arachchige, K. G., Simoes Loureiro, I., Blekic, W., Rossignol, M., and Lefebvre, L. (2021). The role of iconic gestures in speech comprehension: an overview of various methodologies. *Front. Psychol.* 12:634074. doi: 10.3389/fpsyg.2021.634074

Kang, A., Hallman, G. L., Son, L. K., and Black, J. B. (2013). The different benefits from different gestures in understanding a concept. *J. Sci. Educ. Technol.* 22, 825–837. doi: 10.1007/s10956-012-9433-5

Kelly, S. D., Barr, D., Church, R. B., and Lynch, K. (1999). Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory. *J. Mem. Lang.* 40, 577–592. doi: 10.1006/jmla.1999.2634

Kelly, S. D., Creigh, P., and Bartolotti, J. (2010a). Integrating speech and iconic gestures in a Stroop-like task: evidence for automatic processing. *J. Cogn. Neurosci.* 22, 683–694. doi: 10.1162/jocn.2009.21254

Kelly, S. D., Kravitz, C., and Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain Lang.* 89, 253–260. doi: 10.1016/s0093-934x(03)00335-3

Kelly, S. D., Özyürek, A., and Maris, E. (2010b). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychol. Sci.* 21, 260–267. doi: 10.1177/0956797609357327

Kendon, A. (1986). Some reasons for studying gesture. *Semiotica* 62, 3–28. doi: 10.1515/semi.1986.62.1-2.3

Kendon, A. (1997). Gesture. *Annu. Rev. Anthropol.* 26, 109–128. doi: 10.1146/annurev.anthro.26.1.109

Krauss, R. M., Morrel-Samuels, P., and Colasante, C. (1991). Do conversational hand gestures communicate? *J. Pers. Soc. Psychol.* 61, 743–754. doi: 10.1037/0022-3514.61.5.743

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657

Lomber, S. G., Meredith, M. A., and Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat. Neurosci.* 13, 1421–1427. doi: 10.1038/nn.2653

Lyle, S. (2000). Narrative understanding: developing a theoretical context for understanding how children make meaning in classroom settings. *J. Curric. Stud.* 32, 45–63. doi: 10.1080/002202700182844

Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Front. Psychol.* 5:1467. doi: 10.3389/fpsyg.2014.01467

Macedonia, M., Müller, K., and Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Hum. Brain Mapp.* 32, 982–998. doi: 10.1002/hbm.21084

Macedonia, M., and von Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biolinguistics* 6, 393–416. doi: 10.5964/bioling.8931

McNeil, N. M., Alibali, M. W., and Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: now they need it, now they don't. *J. Nonverbal Behav.* 24, 131–150. doi: 10.1023/A:1006657929803

McNeill, D. (1994). Hand and mind: what gestures reveal about thought. *Language* 70, 345–350. doi: 10.2307/415833

Moore, B. C. J. (2014). Development and current status of the “Cambridge” loudness models. *Trends Hear*. doi: 10.1177/2331216514550620

Morett, L. M. (2014). When hands speak louder than words: the role of gesture in the communication, encoding, and recall of words in a novel second language. *Mod. Lang. J.* 98, 834–853. doi: 10.1111/modl.12125

Morett, L. M. (2018). In hand and in mind: effects of gesture production and viewing on second language word learning. *Appl. Psycholinguist.* 39, 355–381. doi: 10.1017/s0142716417000388

Morett, L. M., and Chang, L. Y. (2015). Emphasising sound and meaning: pitch gestures enhance mandarin lexical tone acquisition. *Lang. Cogn. Neurosci.* 30, 347–353. doi: 10.1080/23273798.2014.923105

Morett, L. M., Feiler, J. B., and Getz, L. M. (2022). Elucidating the influences of embodiment and conceptual metaphor on lexical and non-speech tone learning. *Cognition* 222:105014. doi: 10.1016/j.cognition.2022.105014

Moscicki, E. K., Elkins, E. F., Baum, H., and McNarnara, P. M. (1985). Hearing loss in the elderly: an epidemiologic study of the Framingham heart study cohort. *Ear Hear.* 6, 184–190. doi: 10.1097/00003446-198507000-00003

Muir, L. J., and Richardson, I. E. (2005). Perception of sign language and its application to visual communications for deaf people. *J. Deaf. Stud. Deaf. Educ.* 10, 390–401. doi: 10.1093/deafed/eni037

Neville, H. J., and Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: an event-related potential and behavioral study. II. Congenitally deaf adults. *Brain Res.* 405, 268–283. doi: 10.1016/0006-8993(87)90296-4

Obermeier, C., Dolk, T., and Gunter, T. C. (2012). The benefit of gestures during communication: evidence from hearing and hearing-impaired individuals. *Cortex* 48, 857–870. doi: 10.1016/j.cortex.2011.02.007

Obermeier, C., Holle, H., and Gunter, T. C. (2011). What iconic gesture fragments reveal about gesture-speech integration: when synchrony is lost, memory can help. *J. Cogn. Neurosci.* 23, 1648–1663. doi: 10.1162/jocn.2010.21498

Ozyurek, A. (2010). “The role of iconic gestures in production and comprehension of language: evidence from brain and behavior,” in *Gesture in Embodied Communication and Human-Computer Interaction*. eds. S. Kopp and I. Wachsmuth, GW 2009. Lecture Notes in Computer Science, *Vol. 5934* (Berlin, Heidelberg: Springer).

Porter, A. (2016). A helping hand with language learning: teaching French vocabulary with gesture. *Lang. Learn. J.* 44, 236–256. doi: 10.1080/09571736.2012.750681

Proksch, J., and Bavelier, D. (2002). Changes in the spatial distribution of visual attention after early deafness. *J. Cogn. Neurosci.* 14, 687–701. doi: 10.1162/08989290260138591

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532. doi: 10.1037/0033-2909.114.3.510

Reuben, D. B., Walsh, K., Moore, A. A., Damesyn, M., and Greendate, G. A. (1998). Hearing loss in community-dwelling order persons: national prevalence data and identification using simple questions. *J. Am. Geriatr. Soc.* 46, 1008–1011. doi: 10.1111/j.1532-5415.1998.tb02758.x

Rogers, W. T. (1978). The contribution of kinesic illustrators toward the comprehension of verbal behavior within utterances. *Hum. Commun. Res.* 5, 54–62. doi: 10.1111/j.1468-2958.1978.tb00622.x

Sadoski, M. (2018). Reading comprehension is embodied: theoretical and practical considerations. *Educ. Psychol. Rev.* 30, 331–349. doi: 10.1007/s10648-017-9412-8

Schmithorst, V. J., Holland, S. K., and Plante, E. (2006). Cognitive modules utilized for narrative comprehension in children: a functional magnetic resonance imaging study. *NeuroIamge* 29, 254–266. doi: 10.1016/j.neuroimage.2005.07.020

So, W. C., Sim Chen-Hui, C., and Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: is meaning in gesture important for memory recall? *Lang. Cogn. Process.* 27, 665–681. doi: 10.1080/01690965.2011.573220

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Taitelbaum-Swead, R., Brownstein, Z., Muchnik, C., Kishon-Rabin, L., Kronenberg, J., Megirov, L., et al. (2006). Connexin-associated deafness and speech perception outcome of cochlear implantation. *Arch. Otorhinolaryngol.- Head Neck Surg.* 132, 495–500. doi: 10.1001/archotol.132.5.495

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture* 8, 219–235. doi: 10.1075/gest.8.2.06tel

Vendrame, M., Cutica, I., and Bucciarelli, M. (2010). "I see what you mean": oral deaf individuals benefit from speaker's gesturing. *Eur. J. Cogn. Psychol.* 22, 612–639. doi: 10.1080/09541440903126030

Verrecchia, B., and Curcio, V. (2016). Diagnosis of hearing loss in newborns and infants, through objective audiological assessment. *Glob. J. Otolaryngol.* 1, 85–87. doi: 10.19080/gjo.2016.01.555573

Volterra, V., and Erting, C. J. (1994). *From Gesture to Language in Hearing and Deaf Children*. Berlin: Springer-Verlag.

Willems, R. M. (2007). When language meets action: the neural integration of gesture and speech. *Cereb. Cortex* 17, 2322–2333. doi: 10.1093/cercor/bhl141

Willems, R. M., and Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain Lang.* 101, 278–289. doi: 10.1016/j.bandl.2007.03.004

Xi, X., Li, P., Baills, F., and Prieto, P. (2020). Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *J. Speech Lang. Hear. Res.* 63, 3571–3585. doi: 10.1044/2020_jslhr-20-00084

Yuri, A., Platz, E. A., and Niparko, J. K. (2008). Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the national health and nutrition examination survey, 1999-2004. *Arch. Intern. Med.* 168, 1522–1530. doi: 10.1001/archinte.168.14.1522

Zhao, W. Y., Riggs, K., Schindler, I., and Holle, H. (2018). Transcranial magnetic stimulation over left inferior frontal and posterior temporal cortex disrupts gestures-speech integration. *J. Neurosci.* 38, 1891–1900. doi: 10.1523/jneurosci.1748-17.2017

Zhen, A., Van Hedger, S., Heald, S., Goldin-Meadow, S., and Tian, X. (2019). Manual directional gestures facilitate cross-modal perceptual learning. *Cognition* 187, 178–187. doi: 10.1016/j.cognition.2019.03.004

Zheng, A., Hirata, Y., and Kelly, S. D. (2018). Exploring the effects of imitating hand gestures and head nods on L1 and L2 mandarin tone production. *J. Speech Lang. Hear. Res.* 61, 2179–2195. doi: 10.1044/2018_jslhr-s-17-0481

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fpsyg.2022.888226

N400 Indexing the Motion Concept Shared by Music and Words

Tongquan Zhou[1*†], Yulu Li[2*†], Honglei Liu[3], Siruo Zhou[4] and Tao Wang[5]

[1] School of Foreign Languages, Southeast University, Nanjing, China, [2] College of Chinese Language and Literature, Qufu Normal University, Qufu, China, [3] School of Music, Qufu Normal University, Rizhao, China, [4] Department of Chinese Language and Literature, Yonsei University, Seoul, South Korea, [5] School of Psychology, Qufu Normal University, Qufu, China

OPEN ACCESS

Edited by:
Wanjin Meng,
National Institute for Education Sciences, China

Reviewed by:
Linshu Zhou,
Shanghai Normal University, China
Zude Zhu,
Jiangsu Normal University, China

***Correspondence:**
Yulu Li
m15066332357@163.com
Tongquan Zhou
zhoutongquan@126.com

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *02 March 2022*
Accepted: *18 May 2022*
Published: *17 June 2022*

Citation:
Zhou T, Li Y, Liu H, Zhou S and Wang T (2022) N400 Indexing the Motion Concept Shared by Music and Words. Front. Psychol. 13:888226. doi: 10.3389/fpsyg.2022.888226

The two event-related potentials (ERP) studies investigated how verbs and nouns were processed in different music priming conditions in order to reveal whether the motion concept *via* embodiment can be stimulated and evoked across categories. Study 1 (Tasks 1 and 2) tested the processing of verbs (action verbs vs. state verbs) primed by two music types, with tempo changes (accelerating music vs. decelerating music) and without tempo changes (fast music vs. slow music) while Study 2 (Tasks 3 and 4) tested the processing of nouns (animate nouns vs. inanimate nouns) in the same priming condition as adopted in Study 1. During the experiments, participants were required to hear a piece of music prior to judging whether an ensuing word (verb or noun) is semantically congruent with the motion concept conveyed by the music. The results show that in the priming condition of music with tempo changes, state verbs and inanimate nouns elicited larger N400 amplitudes than action verbs and animate nouns, respectively in the anterior regions and anterior to central regions, whereas in the priming condition of music without tempo changes, action verbs elicited larger N400 amplitudes than state verbs and the two categories of nouns revealed no N400 difference, unexpectedly. The interactions between music and words were significant only in Tasks 1, 2, and 3. Taken together, the results demonstrate that firstly, music with tempo changes and music without tempo prime verbs and nouns in different fashions; secondly, action verbs and animate nouns are easier to process than state verbs and inanimate nouns when primed by music with tempo changes due to the shared motion concept across categories; thirdly, bodily experience differentiates between music and words in coding (encoding and decoding) fashion but the motion concept conveyed by the two categories can be subtly extracted on the metaphorical basis, as indicated in the N400 component. Our studies reveal that music tempos can prime different word classes, favoring the notion that embodied motion concept exists across domains and adding evidence to the hypothesis that music and language share the neural mechanism of meaning processing.

Keywords: verbs, nouns, music, motion concept, embodiment

INTRODUCTION

Music and language as the two important communicating systems for human beings are comparable in multiple dimensions (e.g., acoustic features, emotions, and meanings) but mainly in two aspects, syntax and semantics. Syntactically, music and language have similar hierarchical configurations whereby discrete structural elements are combined into sequences (Patel, 2008, p. 241). As an illustration, a section in music is composed of motifs and phrases while a sentence in a language is composed of words and phrases in a hierarchical fashion. This syntactic comparability has been demonstrated in a couple of neuropsychological and neuroimaging studies (e.g., Jentschke et al., 2014; Chiang et al., 2018). Semantically, both music and language are adopted to convey the information that can be interpreted and comprehended by others, despite the point that "the meaning evoked by music is far less specific than meaning evoked by language" (Slevc and Patel, 2011). In Koelsch et al.'s (2004) event-related potential (ERP) study using different types of contexts to prime the processing of words, N400 was elicited by nouns when preceded by either semantically unrelated musical excerpts or semantically unrelated sentences. Afterward, numerous researches converge to reveal the psychological reality of musical meaning similar to linguistic meaning as indexed by N400, in both music-priming-words conditions and words/sentences-priming-music conditions (e.g., Steinbeis and Koelsch, 2008; Daltrozzo and Schön, 2009a,b; Koelsch, 2011). So to speak, meanings are encoded diversely by music and language, but their semantics does partly overlap at least from the perspective of meaning processing.

Musical meaning is abundant in kinds. According to Koelsch (2011), musical meaning can arise from extra-musical sign qualities, intra-musical structural relations, musicogenic effects, the establishment of a unified coherent sense out of "lower-level" units, or musical discourse, which together are generalized into three fundamentally different classes of meaning, extra-musical meaning, intra-musical meaning, and musicogenic meaning. Related to our study is the extra-musical meaning[1] which "emerges from the interpretation of musical information with reference to the extra-musical world", specified as three dimensions—iconic musical meaning, indexical musical meaning, and symbolic musical meaning (Koelsch, 2011). The meaning of musical motion (i.e., motion concept) pertains to iconic musical meaning used to imitate the sounds and qualities of objects or qualities of abstract concepts, in accordance with Eitan and Granot (2006) that music is able to evoke a sense of motion in a listener and with Patel (2008, p. 331) that music can evoke semantic concepts. In light of cognitive linguistics, our understanding of musical motion is completely metaphoric, grounded by our bodily experiences of physical motion (Johnson and Larson, 2003). That is, the motion concept in music can be metaphorically understood *via* embodiment, in fact, a cross-domain mapping from physical motion to musical motion involving our participation (Sloboda, 1998; Todd, 1999; Larson, 2002, 2004; Johnson and Larson, 2003; Eitan and Granot, 2006; Hedger et al., 2013). On the basis of embodiment, the motion concept is encoded not only by music but also by words in the language, for words are the basic categories used to represent an entity, actions, or their relevant features (Wolter et al., 2015). Verbs and nouns as the two major word classes in language are used to represent dynamic objects and static objects separately in general (Shao and Liu, 2001), yet nouns can communicate dynamic characteristics in some ways (as illustrated below). As a result, it is possible to use music to prime verbs and nouns on the shared motion concept basis.

In Mandarin Chinese, verbs have two types of meaning, static meaning and dynamic meaning (Dai et al., 1995), respectively related to state verbs and action verbs. Based on our bodily experience, action verbs and state verbs represent two different subcategories in terms of the motion features on their own. As a consequence, state verbs are often perceived to encode physical state or property and are hence called low-motion verbs, while action verbs, often called high-motion verbs, embrace more motion information than state verbs (Muraki et al., 2020). This point of view is basically consistent with Grossman et al.'s (2002) found that motion-related verbs involve more sensorimotor experience than state verbs, accordingly yielding the relatively easier processing of action verbs. For nouns, animacy is a good indicator for judging their related motion information. According to Weckerly and Kutas (1999), animate nouns as ideal actors in sentences have a strong possibility to perform actions and consequently contain more motion information compared to inanimate nouns. Experiments show that animacy can be clearly distinguished by infants <1 year old based on the motion clues with dynamic or static information (Pauen and Träuble, 2009; Träuble et al., 2014). All these studies suggest that verbs and nouns can convey motion concept but the concept is differently encoded not only between the two-word classes but also between their subcategories.

Motion concept is differently mapped onto music and words. In music, tempo as an expression of extra-musical meaning is used to convey the motion concept metaphorically (Todd, 1999; Johnson and Larson, 2003; Eitan and Granot, 2006; Zhou et al., 2015), involving physical motion and motion imagery. For instance, music tempos with acceleration and deceleration can elicit images of increasingly and decreasingly speeded motion, respectively (Eitan and Granot, 2006; Savaki and Raos, 2019). In the view of Hedger et al. (2013), relative to statistically fast and slow music, accelerating music and decelerating music can better prime pictures in motion and at rest, respectively. This suggests that the music with tempo changes (accelerating music, decelerating music) can better convey motion concept than music without tempo changes (statistically fast and slow music), for the tempo changes in a single piece with acceleration or deceleration may be more apparent and expectable (Hedger et al., 2013). Different from the explicit mapping of motion concept onto music tempo, the motion concept is encoded by verbs and nouns in language implicitly. That is, one's speaking of a verb or

[1] In Koelsch (2011), extral-musical meaning is defined as the interpretation of the musical information associated with external world, which means musical patterns or acoustic properties resembles the emotion state, sounds of objects, qualities of entity, abstract concept and their like. For example, a music excerpt may sound like "a bird"; the accelerating music tempos may sound like "running" movement.

a noun evokes his motion concept covertly and subconsciously. Grounded by this difference, using music tempo to prime verbs and nouns appears more salient than the other way around.

The motivation of music and words able to convey motion concept is well-explained by the embodiment theories and the theory of embodiment semantics. As one of the classical embodiment theories, the perceptual symbols hypothesis (Barsalou, 1999) holds that the symbols are assumed to be the residues of a perceptual experience stored as patterns in the brain for activation. In light of the hypothesis, motion experience is to be activated and simulated as individuals process the motion concept shared by music and words. In music, the acoustic properties of music tempos mimic the properties of physical motion. Similar to the perception of music, the motion concept, which is abstract and implicit in language, is metaphorically encoded by lexical meaning (Hauk et al., 2004; Wolter et al., 2015). To illustrate it, the word –奔跑 "*benpao*/run" is first understood semantically and then its motion attributes as one of its lexical meanings was decoded metaphorically. Also, the motion concept of words is related to embodiment, our sensory-motor experiences (Barsalou, 1999). Likewise, the theory of embodiment semantics claims that the comprehension of lexical meaning is based on our bodily experience and can activate the brain regions responsible for perception, emotion, and action (de Vega et al., 2008; Horchak et al., 2014). This claim has been justified in an fMRI study that sensorimotor experiential traces are activated while processing words referring to an action, e.g., "kicking" can activate the premotor cortex in the brain as actual kicking movement being performed (Hauk et al., 2004).

As stated above, N400 turned out to be an index of semantic incongruity related to extramusical meaning and linguistic meaning. Yet to date, the literature using music to prime words' meaning or motion concept has been confined to three studies by Koelsch et al. (2004), Hedger et al. (2013), and Zhou et al. (2015). Hedger et al. (2013) conducted a behavioral study to reveal that accelerating music and decelerating music can better prime motion concept than fast music and slow music and the incongruency motion relations between tempos and pictures cost more time than the congruency. In Koelsch et al. (2004)'s ERP study, music excerpts were proved to be as valid as sentences to facilitate semantically congruent words, as revealed by the N400 in both music and sentence conditions. Similarly, Zhou et al. (2015) drew up music excerpts to prime semantically congruent and incongruent pictures in a set of ERP experiments to indicate that incongruent pairs elicited a larger N400 than the congruent pairs over the anterior and central regions, further justifying the role of N400 in revealing the motion concept conveyed by music. The three studies converge to show that music can convey motion concept and other meanings related to words or pictures. Nevertheless, scrutiny of their experiments uncovers some gaps to be filled: for one thing, word types as the target stimuli were not rigidly manipulated–e.g., concrete nouns were confused with abstract nouns, and mono-category words confused with multi-category words in Koelsch et al. (2004) and the heterogeneity of stimuli may fail to reveal the priming effect as anticipated; for the other, the motion concept-based (in)congruency was established between music and pictures as in Hedger et al. (2013) and Zhou et al. (2015), leaving an open question whether music can exert influence on verbal stimuli like verbs and nouns so that motion concept as a putative shared meaning can be evoked across more domains.

Against the above background, the current study referring to Hedger et al. (2013) and Zhou et al. (2015) utilizes two ERP experiments (Studies 1 and 2) to explore (1) whether music tempos can prime Chinese verbs and nouns, (2) how the congruency between music and words is established on their shared motion concept basis, and (3) how the four classes of verbs and nouns are distinguished from the perspective of processing under music priming conditions. Specifically, Study 1 tested how the two types of music (with vs. without tempo changes) affect the processing of two sub-classes of verbs (action verbs vs. state verbs) while Study 2 tested how the same two types of music affected the processing of two sub-classes of nouns (animate nouns vs. inanimate nouns). In association with previous relevant research, we make the predictions as below:

First, music with tempo changes could better facilitate the processing of verbs and nouns than music without tempo changes, yielding a reduced N400 effect by the words;

Second, on the basis of the shared motion concept, the incongruent pairs relative to the congruent pairs between music and words may elicit larger N400 amplitudes in the words;

Third, action verbs and animate nouns should be easier to process than state verbs and inanimate nouns, respectively. This relative ease was due to individuals' sensorimotor experience with more motion information related to the two subclasses of words.

STUDY 1

As a rule, verbs are used to typically represent actions, and music is characterized by tempo changes, a way of motion representation. Based on this conceptual similarity, the first study investigated whether and how music could activate the processing of verbs by virtue of the cross-modality concept priming paradigm. Two types of music (with or without tempo changes) were selected to prime two types of Chinese verbs (action verbs or state verbs). The whole experiment was composed of two tasks by a within-subjects design, in which Task 1 adopted the music with tempo changes (accelerating music vs. decelerating music) and Task 2 with the music without tempo changes (fast music vs. slow music) to prime verbs (action verbs vs. state verbs).

In light of the above-mentioned music type and verb type, four pairs of priming conditions were designed for Task 1 as shown in **Figure 1**: accelerating music—action verb pair, decelerating music—state verb pair to constitute a congruent relation whereas decelerating music—action verb pair and accelerating music—state verb pair to constitute an incongruent relation. Task 2 was similar to Task 1, with the only difference that the prime stimuli were music without tempo changes (fast music or slow music). Specifically, fast music—action verb pair and slow music—state verb pair constituted a congruent relation whereas slow music—action verb pair and fast—state verb pair constituted an incongruent relation. Such a design was referred to Hedger et al.'s

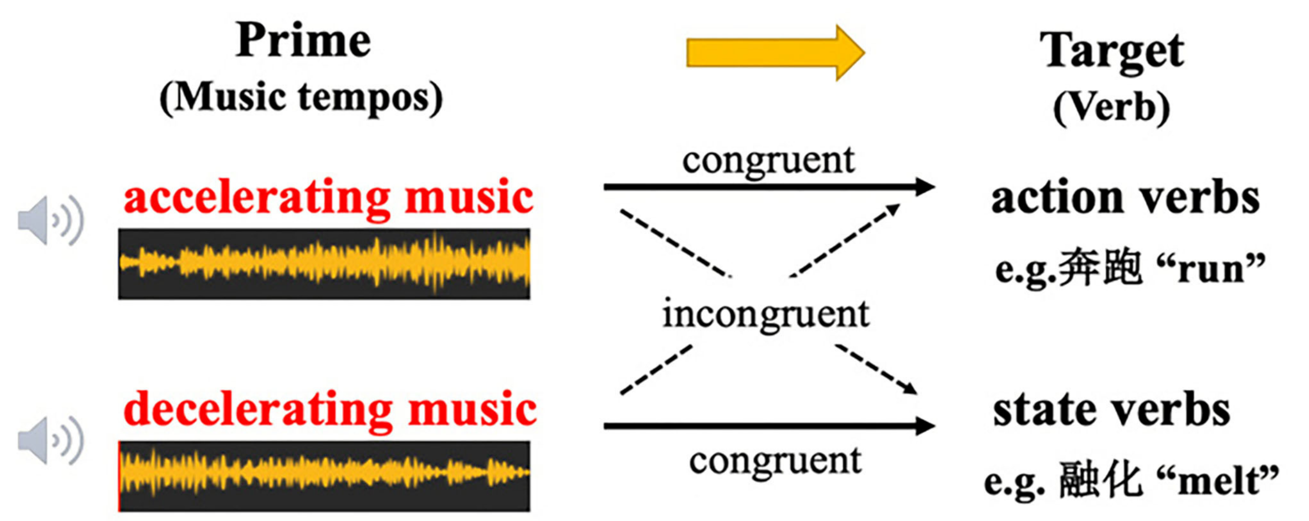

FIGURE 1 | Design of the cross-modality priming paradigm in task 1 of Experiment 1. Music tempos were used as primes, and verbs were used as targets.

(2013) finding that accelerating music and decelerating music could better prime pictures in motion and at rest separately, and in our experiment action verbs and state verbs were adopted to replace the pictures in motion and at rest so as to build up congruent and incongruent stimulus pairs.

Methods

Participants

A total of 40 Qufu Normal University students (age: M = 20.1 years, SD = 1.5 years, ranging from 18 to 24 years of age; gender: 36 women, four men) were recruited to participate in the experiment as paid volunteers. All the participants are native Chinese speakers, right-handed in terms of the Edinburgh Handedness Inventory (Oldfield, 1971), with normal hearing and normal or corrected to normal vision and no history of psychiatric or neurological diseases. They have no musical experience (none had received professional musical training or played any instruments). All participants signed a formal written consent before the experiment. The experiment was approved by the Ethics Committee of Qufu Normal University. The data of two participants were discarded due to excessive drift artifacts during the experiment. Therefore, the data to enter into *post-hoc* analysis consisted of 38 participants (age: $M = 20.03$ years, SD = 1.46 years; gender: 34 women, four men).

Stimuli Construction

Priming stimuli were music motifs. In Task 1, 20 music motifs were created as the priming stimuli, which were subdivided into 10 accelerating music motifs and ten decelerating music motifs and reduplicated twice for each. In Task 2, the priming music was changed into the one without tempo-change motifs (fast music and slow music). All the music motifs were created by a MIDI controller and audio editor 3.2.9, 44 kHz, 16-bit resolution, with an average duration of 10 s. These music stimuli consisted of two oscillating notes which were alternatively processed to form a rhythm (see Hedger et al., 2013 for similar manipulations). Participants listened to the music *via* Sony MRD-XB55AP headphones prior to making (in) congruency judgment between the music and the visualized words on the computer screen. In Task 1, accelerating music (which began with a tempo of 120 BPM and ended with 600 BPM) represented the music with strong motion information while decelerating music (from 600 to 120 BPM) represented the music with static motion information. In Task 2, fast music motifs (at 600 BPM from the beginning of a note to its end) and slow music motifs (at 120 BPM) are linked with music with strong motion and static motion information, respectively. Additionally, 20 music motifs with irregular tempo changes were selected as fillers for each task.

The target stimuli were Chinese verbs, comprised of 20 action verbs and 20 state verbs in each task, yielding 40 target stimuli for each task. No verbs in Task 1 re-occurred in Task 2. Target words were taken from the CCL corpus (Center for Chinese Linguistics, Peking University) with high frequency. According to Hu et al. (1989), action verbs were selected by referring to the following criteria: (1) the verbs signal action other than state; (2) the action is autonomous; (3) the action is concrete other than virtual or abstract; (4) each action doer (agent) has the strong executive ability and individual motivation. In accordance with the definition of state verb by different scholars (Yuan, 1998; Chen, 2002; Ma, 2005), forty state verbs were selected by referring to the following criteria: (1) the verbs are non-autonomous; (2) the verbs signal sustainable state or property other than action; (3) the verbs are not used perfectively; (4) verbs are non-bodily and have no spatial displacement. The stoke number of words were balanced (action verb: $M = 17.3$, $SD = 4.04$; state verb: $M = 18.7$, $SD = 4.071$), with independent t-test showing that they were not systematically different [$t_{(78)} = -1.544$, $p = 0.127$]. A total of 20 conjunctions were selected as fillers for each task.

Normalization of Materials

In order to obtain the stimuli for experiments, we normalized the materials by testing the motion attributes and imaginability of verbs and the relatedness between music and verbs prior to the experiment.

In the motion attributes test and imaginability test, action verbs and state verbs were selected to represent high-motion words and low-motion words, respectively, and all verbs with high imaginability. Before the experiment, 60 action verbs and 60 state verbs on a seven-point scale were tested by 100 participants who did not participate in the formal experiment, ranging from −3 (very low in motion) to +3 (very high in motion) the test motion attribute and −3 (very low imaginability) to +3 (very high imaginability) for the test of imaginability. Finally, 40 action verbs and 40 state verbs were selected as experimental materials. The average score of action verbs with high motion information evaluation was 2.165 ($SD = 0.213$) and state verbs with low motion information evaluation was −2.105 ($SD = 0.214$). In terms of imaginability, the average scores of action verbs and state verbs were 2.045 ($SD = 0.281$) and 2.016 ($SD = 0.199$). An independent samples t-test showed that the two classes of verbs were significantly different from each other with regard to motion information [$t_{(78)} = 89.483$, $p < 0.01$] but not significantly different in imaginability [$t_{(78)} = 0.523$, $p = 0.603$].

In order to examine whether music is related to verbs with respect to the motion concept, we tested their congruency relations on the basis of a seven-point scale, from −3 (strong unrelated) to +3 (strong related), with 0 signaling uncertainty. Only the excerpts with a mean rating score of higher than +1 or lower than −1 were selected as experimental materials. Results showed that the mean scores of the congruent groups in Tasks 1 and 2 were 2.011 ($SD = 0.246$) and 1.876 ($SD = 0.246$), respectively, and the mean rating scores of the incongruent groups in Tasks 1 and 2 were −1.966 ($SD = 0.271$) and −1.734 ($SD = 0.182$), respectively, suggesting that people can well-establish the congruency between music and nouns perceptively. Besides, t-test revealed that the congruent pairs and the incongruent pairs were significantly different in both Task 1 [$t_{(38)} = 48.664$, $p < 0.01$] Task 2 [$t_{(38)} = 52.819$, $p < 0.01$]. According to the rating results, 40 pairs in each task (10 pairs in each condition) were chosen as final experiment materials for subsequent experiments. Items and conditions were counterbalanced in each task. All the materials were presented in random order and designed *via* E-prime 3.0 (Psychology Software Tools, Inc.).

Procedures

Participants were seated on a comfortable chair in front of the computer screen approximately one meter away in a soundproof

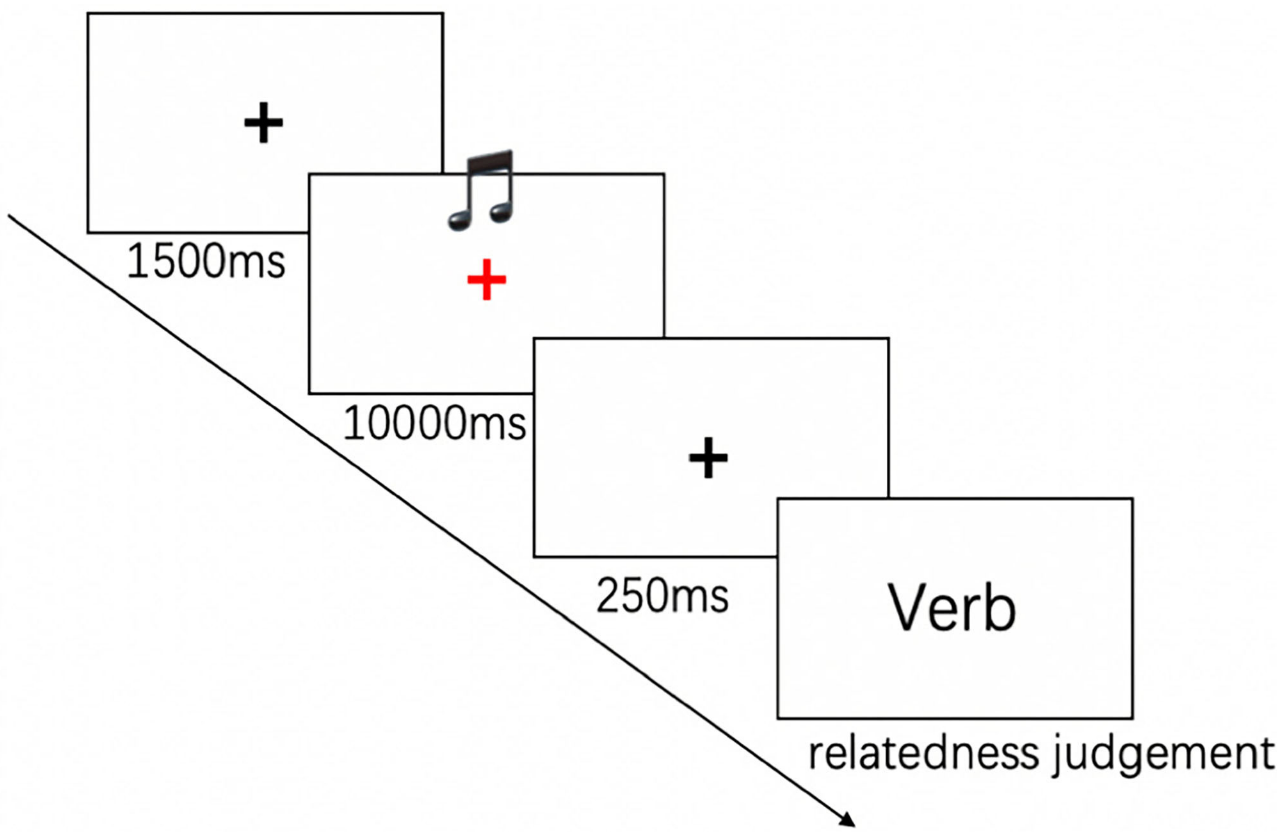

FIGURE 2 | The procedure of Experiment 1.

room. In each trial, the participants heard a piece of music *via* the earphone they were wearing, followed by a verb at the center of the screen on a computer. Their tasks were to as quickly as possible judge whether or not the verb was congruent with the music motif in terms of motion information (pressing the button "F" for congruent or "J" for incongruent), with the procedure as shown in **Figure 2**. As the experiment started, there appeared a black fixation cross (lasting 1,500 ms) at the center of the screen against a white background. As the cross turned red, the music as the priming stimulus was broadcasted for 10 s. Afterward, a second black fixation cross came out as an interval for 250 ms until each participant made the congruency judgment on the target verb. No response time limitation was set for the judgment on the verb. All participants were required not to blink their eyes as much as possible with the exception of the interval (1,500 ms black fixation cross) between each trail. The order of all the trails was presented randomly, with each trail for once. The entire experiment lasted ∼14 min.

EEG Recording and Preprocessing

The EEG data were recorded by AC amplifiers from 64 scalp locations of the International 10–20 system. The electrodes placed on the left and right mastoids served as the reference, and the electrode between Fz and Cz was selected as the ground. To filter the eye movement and eye blink, the data of the horizontal electrodes placed on the canthus of each eye and vertical electrodes above and below the left eye were recorded. Impedances of the electrodes were kept below 5 kΩ. EEG data were digitized with the rate of 1,000 Hz and amplified and filtered within a band-pass of 0.1 and 30 Hz. Trails with artifacts (eye movements, head movements) exceeding the amplitude of ±200 μV at any channel were excluded. Raw EEG data were preprocessed by NeuroScan SynAmps2 8050 (Compumedics Neuroscan) and Curry 8 software (Compumedics Neuroscan). Further processing was carried out by EEGLAB 14.1.1 (Delorme and Makeig, 2004) in MATLAB 2013b (MathWorks).

ERP Data Analysis

All the data were analyzed by computing the mean amplitudes under each condition for each task. ERP data were segmented from 200 ms before to 800 ms after the onset of the target, with a 200 ms pre-stimulus correct baseline. Based on the previous studies on the N400, a time window from 300 to 500 ms after target-stimulus onset was chosen for statistical analysis.

Preliminary analysis of congruency level did not show the processing advantages of congruent pairs or incongruent pairs, suggesting the relations between music and words are more complex. Given that different verbs have different attributes and processing patterns (e.g., Amsel and Cree, 2013; Muraki et al., 2020: nonbodily state abstract verbs can elicit a larger N400 component than concrete verbs), we referred to Hedger et al. (2013) to reanalyze the data by selecting the music tempos and verbs as conditions so as to observe the processing of each type of word. The result shows that the motion concept indexed by N400 can be revealed by different types of music and words but not by their congruency levels.

To test the distribution of the effects, nine-item regions of interest (ROIs) were selected, with details shown in **Table 1**. For

TABLE 1 | Electrode channels in each region of interest (ROI).

ROI	Left	Central	Right
Anterior	FP1, AF3, F1, F3, F5, F7	FPz, Fz	FP2, AF4, F2, F4, F6, F8
Central	FC3, FC5, C3, C5, CP3, CP5	FCz, Cz, CPz	FC4, FC6, C4, C6, CP4, CP6
Posterior	P3, P5, P7, PO3, PO5, PO7	Pz, POz, Oz	P4, P6, P8, PO4, PO6, PO8

statistical analysis, ERPs were analyzed by repeated-measures ANOVAs to test the effects among music, verb, hemisphere, and regions. Nine regions of interest were divided into lateral electrode regions and midline electrode regions. The mean amplitudes of the lateral and midline electrode regions were computed separately for analysis. For the lateral electrodes, music (accelerating music and decelerating music in Task 1, fast music and slow music in Task 2), verb (action verbs and state verbs), region (anterior, central, and posterior) and hemisphere (left hemisphere and right hemisphere) served as within-subjects factors. For the midline electrodes, music (accelerating music and decelerating music in Task 1, fast music and slow music in Task 2), verb (action verbs and state verbs), and regions (anterior, central, and posterior) served as within-item factors.

In the analysis of the general linear model, data were adjusted by the Bonferroni correction. All p-values reported below were adjusted with the Greenhouse–Geisser correction when the degree of freedom in the numerator was larger than 1. The reported eta squared (η^2) was used to measure the effect size for ANOVAs (Olejnik and Algina, 2003).

Results

Behavioral Results

Behavioral results were analyzed in R studio 3.5.1 software (R Development Core Team, 2012). Given the poor performance of participants would influence the results, the data were deleted for the trials in which the RTs were shorter than 500 ms or longer than 3SD above the overall average. As a result, 9.2 and 7.68% of the collected data were discarded in Tasks 1 and 2, respectively. The accuracy (ACC) and reaction time (RT) were recorded by computer automatically. In Tasks 1 and 2, the participants responded with a mean accuracy of 91.55% ($SD = 0.258$) and 92.45% ($SD = 0.25$), respectively, indicating that they followed the instructions and attended to the stimuli carefully. The mean reaction time in Tasks 1 and 2 was 1,136.75 ms ($SD = 339.5$) and 1,137.75 ms ($SD = 36$), respectively. A main effect of verb type was found in both tasks (Task 1: $\beta = 0{,}112$, $SE = 0.041$, $df = 70.275$, $t = 2.73$, $p < 0.01$; Task 2: $\beta = 0.155$, $SE = 0.036$, $df = 39.450$, $t = 4.361$, $p < 0.01$), i.e., the participants responded faster in action verbs condition than in state verbs condition.

Electrophysiological Results

The grand averaged ERP waveform elicited by actions verbs and state verbs in different music conditions are shown in **Figure 3** (Task 1) and **Figure 4** (Task 2) respectively.

In Task 1, the ERP waves to action verbs and state verbs diverged at ∼250 ms after the target verb onset and reached maximality at 400 ms and afterward lasted one more 100 ms (at

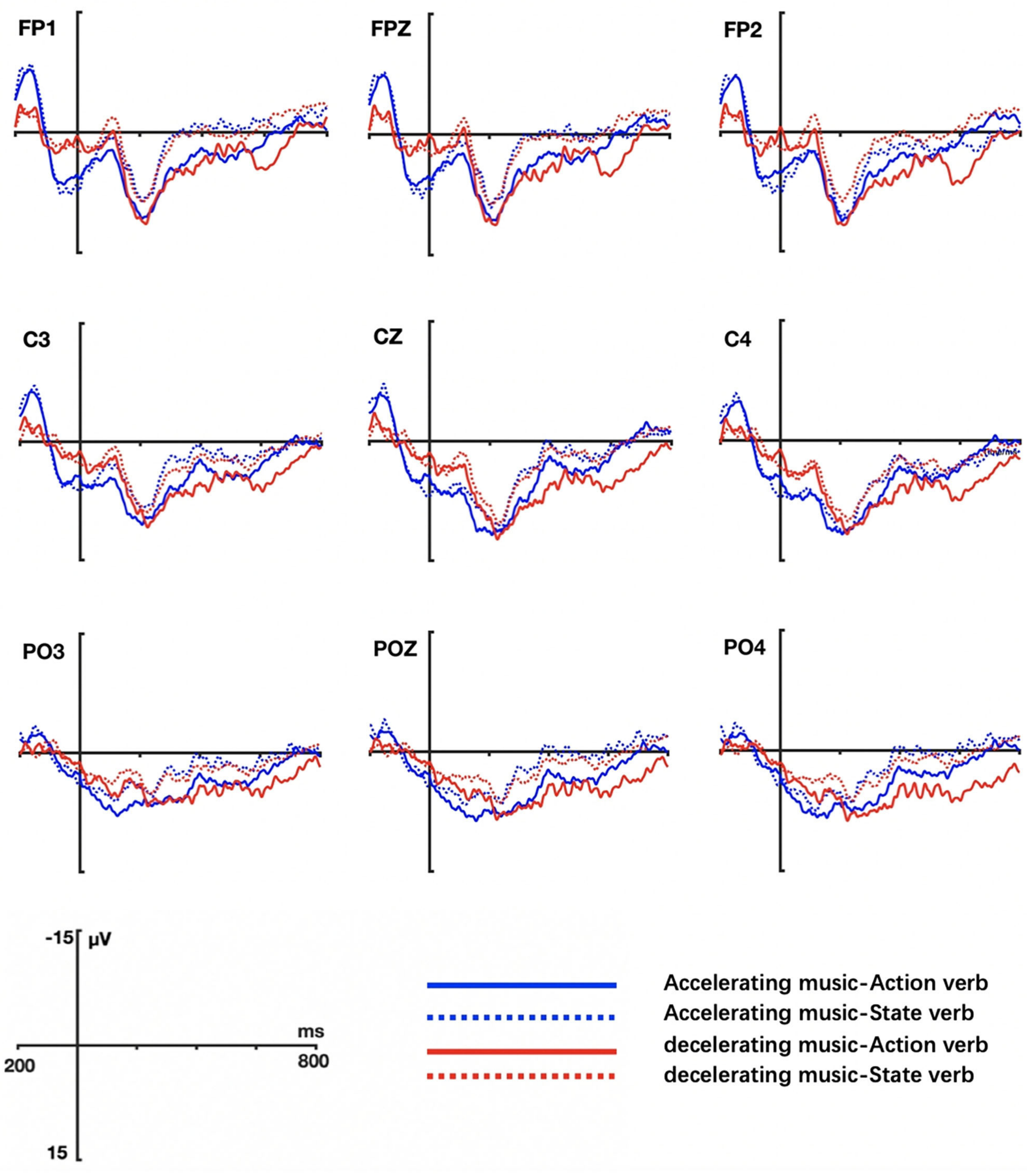

FIGURE 3 | Grand ERP averages elicited by action and state verbs in tempo-with-change music priming conditions (Task 1).

500 ms) (as shown in **Figure 3**). Statistical analysis showed that there was a significant main effect of verb (lateral electrodes: $F = 7.557$, $p = 0.009$, $\eta^2 = 0.17$) in that state verbs elicited larger N400 amplitudes than action verbs independent of music condition. In addition, there occurred a significant interaction between music and verb type (middle electrodes: $F = 10.593$, $p = 0.002$, $\eta^2 = 0.223$). Further simple effect test on verbs showed that state verbs compared to action verbs evoked enhanced

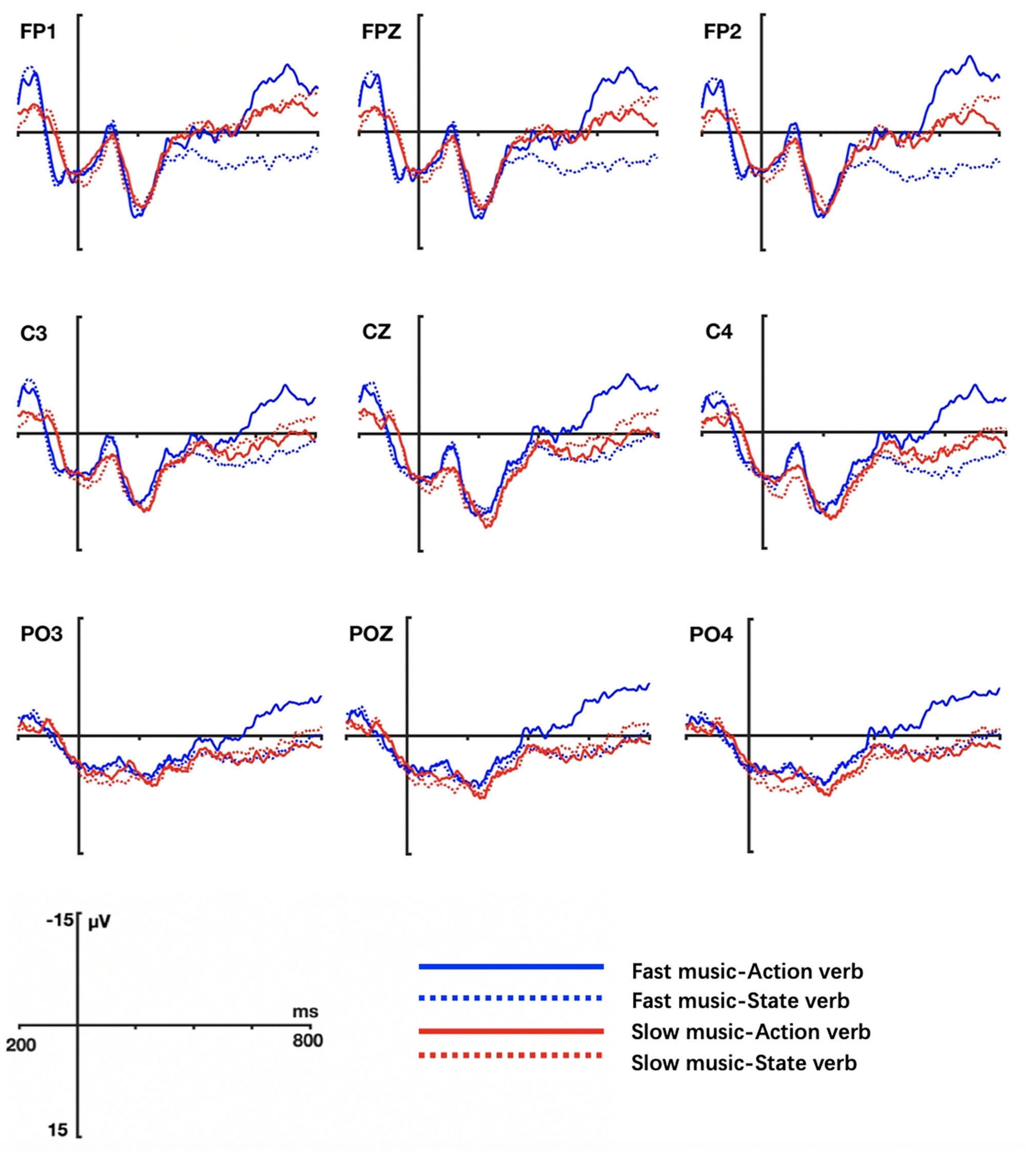

FIGURE 4 | Grand ERP averages elicited by action and state verbs in tempo-without-change music priming conditions (Task 2).

N400 amplitudes in both accelerating music (lateral electrodes: $F = 9.309$, $p = 0.004$, $\eta^2 = 0.201$) and decelerating music conditions (middle electrodes: $F = 10.591$, $p = 0.002$, $\eta^2 = 0.223$). Another simple effect test on music revealed that action verbs induced larger amplitudes in accelerating music condition than in decelerating music condition around the anterior region (middle electrodes: $F = 5.108$, $p = 0.03$, $\eta^2 = 0.121$) (as exhibited in **Figure 5**). Significant interactions were observed between verb and hemisphere ($F = 4.132$, $p = 0.049$, $\eta^2 = 0.100$) but no significant interactions were found between music and

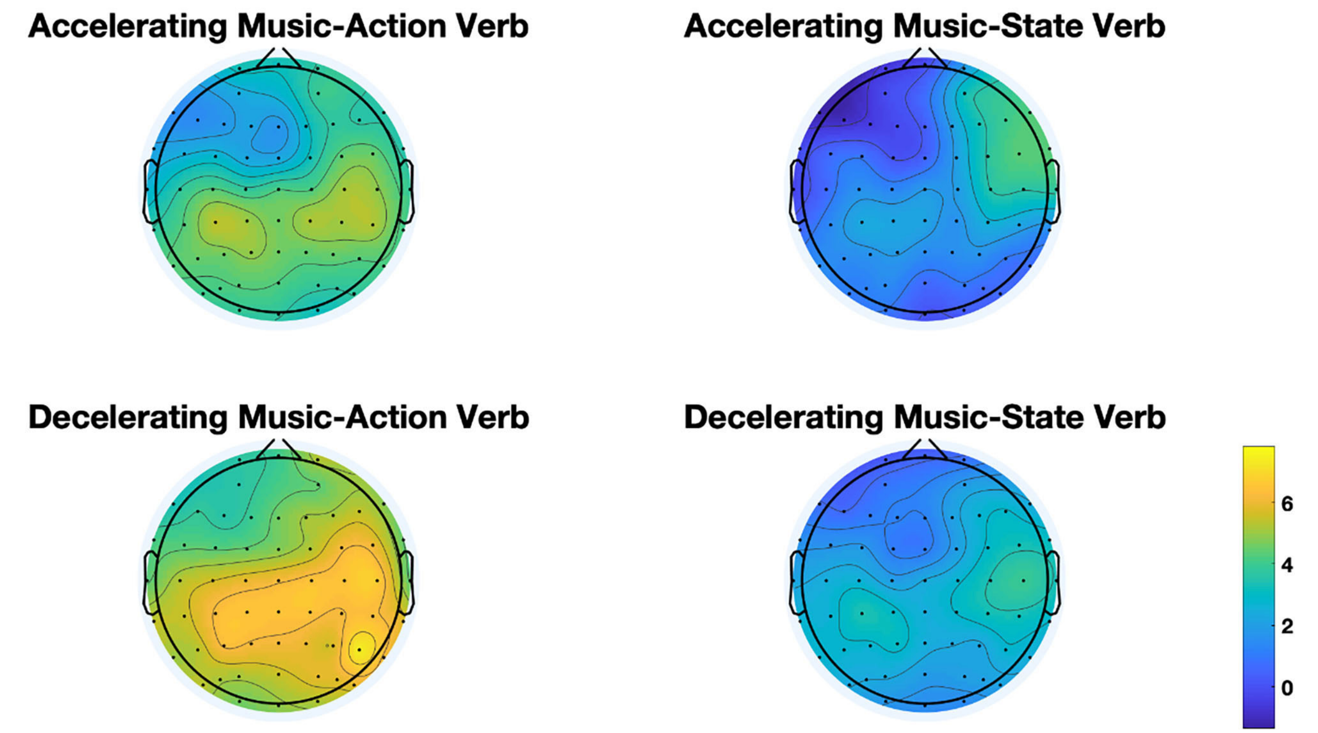

FIGURE 5 | Scalp topographic maps of N400 across four conditions in Task 1.

hemisphere ($F = 0.16$, $p = 0.692$, $\eta^2 = 0.004$), music and region (lateral: $F = 2.495$, $p = 0.119$, $\eta^2 = 0.063$; middle: $F = 1.347$, $p = 0.252$, $\eta^2 = 0.036$), and between verb and region factor (lateral: $F = 2.178$, $p = 0.142$, $\eta^2 = 0.056$; middle: $F = 2.349$, $p = 0.132$, $\eta^2 = 0.060$). Moreover, there were not three-way interactions among music, verb, hemisphere ($F = 2.455$, $p = 0.126$, $\eta^2 = 0.062$) and music, verb, region factors (lateral: $F = 2.518$, $p = 0.112$, $\eta^2 = 0.064$; middle: $F = 2.506$, $p = 0.117$, $\eta^2 = 0.063$).

To sum up, electrophysiological results of task 1 showed that state verbs elicited a larger N400 than action verbs, independent of music type over the anterior regions at the array of analyzed electrodes. Moreover, action verbs relative to state verbs were more sensitive to music type, i.e., action verbs elicited a larger amplitude in accelerating music conditions than in decelerating music conditions during processing action verbs.

In Task 2, the results were quite different. Statistical analysis showed that the main effect of verb type was not significant at both lateral electrodes ($F = 0.018$, $p = 0.893$, $\eta^2 = 0.000$) and middle electrodes ($F = 0.045$, $p = 0.833$, $\eta^2 = 0.001$), indicating that action verbs and state verbs were not so differently processed when primed by the music without tempo changes (fast or slow music). However, there was a significant interaction between music type and verb in the midline ($F = 7.985$, $p = 0.008$, $\eta^2 = 0.177$) and lateral electrodes regions ($F = 9.22$, $p = 0.004$, $\eta^2 = 0.199$). The simple effect test on verb indicated that action verbs elicited larger N400 amplitudes than state verbs in both fast music condition (lateral electrodes: $F = 6.422$, $p = 0.016$, $\eta^2 = 0.148$) and slow music condition (lateral electrodes: $F = 5.887$, $p = 0.02$, $\eta^2 = 0.137$; middle electrodes: $F = 5.682$, $p = 0.022$, $\eta^2 = 0.133$). There were no significant interactions between music and electrode position factors and between verb and electrode position factors, both for lateral (music × hemisphere: $F = 0.101$, $p = 0.752$, $\eta^2 = 0.003$; verb × hemisphere: $F = 1.650$, $p = 0.687$, $\eta^2 = 0.004$; music × region: $F = 0.237$, $p = 0.643$, $\eta^2 = 0.006$; verb × region: $F = 0.749$, $p = 0.418$, $\eta^2 = 0.02$) and middle electrodes (music × region: $F = 0.749$, $p = 0.418$, $\eta^2 = 0.02$; verb × region: $F = 0.072$, $p = 0.889$, $\eta^2 = 0.002$). No three-way interactions among music, verb, hemisphere (lateral: $F = 3.712$, $p = 0.062$, $\eta^2 = 0.091$) and music, verb, region factors (lateral: $F = 0.323$, $p = 0.594$, $\eta^2 = 0.009$; middle: $F = 0.415$, $p = 0.662$, $\eta^2 = 0.011$) were found.

In brief, the music without tempo changes affected the processing of verbs. Contrary to the situation in Task 1, action verbs induced a larger N400 than state verbs in both music conditions.

Discussion

Study 1 intended to investigate how Chinese verbs were processed in different music priming conditions provided that verbs and music take on the shared motion concepts. The ERP data revealed that the effects of different music tempo types were varied. The N400 was elicited by state verbs in Task 1 but by action verbs in Task 2, implying that music with tempo changes facilitated the processing of action verbs but music without tempo changes inhibited the processing of action verbs. That is, change-in-tempo music was more related to the motion concept and triggered the cognition of action verbs as a result. This result partly justifies the first prediction that music with tempo changes can facilitate the processing of verbs.

Unexpectedly, the motion congruency effect by N400 was not found in the two tasks. In different music conditions, action verbs in music with strong motion information conditions (accelerating music in Task 1 and fast music in Task 2) induced larger ERP amplitudes, contrary to our second prediction that processing congruent pairs (accelerating music-action verbs; fast music-action verbs) should be easier. This result was somewhat comparable to the word-repetition situation in which the resulting ERP wave of a word becomes stronger when it is primed on its own than by a different but relevant word (Michael, 1985). Here, we may assume that it is the repetition of motion concepts shared by music and action verb that leads to the increased ERP amplitudes relating to the processing of action verbs. Another related possibility is that establishing the congruency between music and verbs is similar to the music-picture pairs manipulations in Hedger et al. (2013) but more difficult due to the inherent greater complexity in words than in pictures, for the motion concept is encoded directly by pictures but indirectly by words.

The results also demonstrate that the processing patterns between action verbs and state verbs vary with music types. The attenuated N400 amplitudes of action verbs suggest that the processing of action verbs is easier and compatible with the third prediction. In addition, the significant difference between music conditions observed only in action verbs suggests that the sub-types of music with tempo changes only affect action verbs but not state verbs. That is, action verbs seem more sensitive to music conditions. In our study, state verbs elicited enhanced N400 amplitudes, confirming Muraki et al.'s (2020) finding that non-bodily state verbs elicited larger N400 amplitudes than concrete verbs in a syntactic classification task. The motion information in action verbs appears to be simulated more easily than in state verbs due to the more motion information they share.

In brief, the different effects of music tempo types on different verbs verify the idea that people tend to analogically understand the motion concept in music and language by virtue of the acoustic features and linguistic meaning. In this process, bodily motion experience is supposed to be activated as soon as the music and the verbs with high motion information are presented. Yet this experience does not bring about the absolutely similar activation in music and words because they are two different categories and hence distinctly encoded with regard to motion concept.

STUDY 2

Verbs and nouns are two important word classes in Chinese. As revealed in Study 1, verbs have motion attributes, but in folk cognition, verbs are assumed to contain more dynamic information while the majority of nouns contain more static information. Neurophysiological studies show that verbs and nouns are differently processed (e.g., Pulvermüller et al., 1999). In addition, nouns (just like verbs) can be subdivided into many subcategories, among which is the dichotomy of animate nouns and inanimate nouns in terms of the trait (±ANIMACY). In linguistics, it is generally agreed that animate nouns are more associated with action than inanimate nouns, for the former is used to signify objects in motion whereas the latter signals entities at rest in general (Weckerly and Kutas, 1999). On this account, Study 2 was conducted to compare how to animate nouns and inanimate nouns were processed in different priming-music conditions, and the music pieces were just the same sets as adopted in Study 1.

Methods

Just like Study 1, Study 2 adopted the cross-modality concept priming paradigm by using music to prime nouns. The experiment was comprised of two tasks, i.e., Task 3 testing nouns primed by music with tempo changes (accelerating music; decelerating music) and Task 4 testing the nouns primed by music without tempo changes (fast music; slow music). All the targets were animate and inanimate nouns, with half for each. The priming stimuli and target stimuli constituted four congruency conditions, with two congruent pairs (accelerating music- animate nouns, decelerating music-inanimate nouns) and two incongruent pairs (accelerating music-inanimate nouns, decelerating music-inanimate nouns). The whole experiment lasted around 14 min.

Participants

The participants were those who had participated in Study 1. After the experiment, the data of one participant were discarded due to the EEG equipment malfunctioning. Another participant's data were excluded due to her failure to follow the experiment instructions. As a result, there thirty-eight participants' data were kept for analysis. All the participants signed the consent form prior to the experiment and were paid for their participation at the terminal of the experiment.

Stimuli Construction

The priming stimuli of Study 2 were identical to those adopted in Study 1 and the target stimuli were replaced by nouns (from verbs). Priming stimuli in Tasks 3 (accelerating music and decelerating music) and 4 (fast music and slow music) repeated the music motifs in Tasks 1 and 2, respectively, including ten music motifs in each condition. Forty animate nouns and forty inanimate nouns selected from the CCL corpus were used as target stimuli in the two tasks. Each task involved 20 animate nouns and 20 inanimate nouns. No nouns were repeated in the two tasks. All the nouns were concrete and with high frequency, excluding emotional words. The animate nouns were featured by agentivity and high probability in motion or action. For example, the noun 豹子 "*baozi*/ leopard" has the ability to perform an action (e.g., running) and the abstract motion concept is encoded in this concrete entity. By contrast, inanimate nouns were characterized by null agentivity and high staticity, such as 沙发 "*shafa*/sofa" and 雕像 "*diaoxiang*/statue", which often serve as the undergoer to receive action(s).

Stimuli Pre-tests

As in Study 1, in order to manipulate the motion concept of nouns as well as their relatedness with music tempos, we made the pre-tests on the basis of seven-point scales. The steps were identical to those in Study 1.

In the motion attribute test, the mean score of the animate nouns was 1.952 ($SD = 0.248$) and the inanimate nouns −2.114 ($SD = 0.245$), suggesting that the participants could clearly distinguish animate nouns with motion concept from inanimate nouns without motion concept. An independent sample *t*-test showed that the two classes of nouns were significantly different from each other with regard to motion information [$t_{(78)} = 73.734, p < 0.01$].

In the imaginability test, the mean scores of the animate nouns and inanimate nouns were 2.252 ($SD = 0.241$) and 2.197 ($SD = 0.198$), respectively. The independent *t*-test showed that there was no significant difference in their imaginability [$t_{(78)} = 1.114, p = 0.269$].

In the relatedness test, four conditions (pairs) and two relatedness relations (congruent or incongruent) were formed in Study 2, comprising Tasks 3 and 4. Congruent pairs were those with congruent motion information: accelerating (or fast) music—animate noun pairs; decelerating (or slow) music—inanimate noun pairs. On the opposite, incongruent pairs were those with incongruent motion information: accelerating (or fast) music—inanimate noun pairs; decelerating (or slow) music—animate noun pairs. Results showed that the mean scores of the congruent pairs in Tasks 3 and 4 were 1.967 ($SD = 0.226$) and 2.010 ($SD = 0.248$) respectively whereas the mean score of the incongruent pairs in Tasks 3 and 4 are −1.880 ($SD = 0.201$) and −1.708 ($SD = 0.191$), respectively, suggesting that most adults can perceptively (i.e., *via* auditory and visual embodiment) establish the relations between music and nouns. Further comparison showed that the related pairs were strikingly different from the unrelated pairs in both Tasks 3 [$t_{(38)} = 56.939, p < 0.01$] and 4 [$t_{(38)} = 53.101, p < 0.01$]. The two classes of nouns' stroke number were manipulated (animate noun: $M = 17.700$, $SD = 4.767$; inanimate noun: $M = 16.875$, $SD = 4.598$), with the independent *t*-test showing that they were not significantly different [$t_{(78)} = 0.788, p = 0.433$].

According to the rating results, 40 pairs in each task (with 10 pairs in each condition) were chosen as the formal stimuli for the experiment afterward. Items and conditions were counterbalanced in each task, controlled *via* E-prime 3.

Procedures

The procedure was identical to Study 1.

EEG Recordings and Preprocessing

The recording methodologies and the data preprocessing were identical to those in Study 1.

ERP Data Analysis

The data analysis of Study 2 followed the same procedure as in Study 1.

Results

Behavioral Results

To follow the general norm for experimental studies, we deleted the data with RTs shorter than 500 ms or longer than 3 SD above the overall average. As a result, 3.82 and 4% of the collected data were discarded in Tasks 3 and 4, respectively. Behavioral results showed that the participants responded with a mean accuracy of 94.44% ($SD = 0.211$) and 93.8% ($SD = 0.181$) for the two tasks, respectively, indicating that they followed the instructions and attended to the stimuli carefully. The mean reaction time of Tasks 3 and 4 was 1,019.75 ms ($SD = 344.25$) and 1,051.25 ms ($SD = 322.25$), respectively. A significant main effect of noun were observed in the two tasks, in which participants had a faster response to animate nouns than to inanimate nouns (Task 3: $\beta = 0.101$, $SE = 0.027$, $df = 28.043$, $t = 3.726$, $p = 0.001$; Task 4: $\beta = 0.119$, $SE = 0.023$, $df = 31.83$, $t = 5.108$, $p < 0.01$), suggesting animate nouns seem to be easier to process than inanimate nouns.

Electrophysiological Results

The grand averaged ERP waveform elicited by animate nouns and inanimate nouns in different music conditions were shown in **Figure 6** (Task 3) and **Figure 7** (Task 4).

The results of Task 3 were comparable to those of Task 1 in Study 1. Visual inspection showed that larger negativity amplitudes were induced by inanimate nouns, which started from 250 ms and peaked at 400 ms, and lasted to 500 ms (as illustrated in **Figure 6**). ANOVA was utilized to test the ERPs of a 2 music (Task 1: accelerating music, decelerating music; Task 2: fast music, slow music) × 2 noun (animate noun, inanimate noun) × 2 hemispheres (left hemisphere, right hemisphere) × 3 regions (anterior, central, posterior) in this experiment.

In the noun conditions, the analysis showed an observable effect of noun only at lateral electrodes ($F = 17.413$, $p = 0.000$, $\eta^2 = 0.32$). To be precise, compared with animate nouns, inanimate nouns incurred increased N400 amplitudes over the anterior and central regions (as exhibited in **Figure 8**), similar to Zhou et al.'s study (2015), in which the target stimuli were non-verbal pictures. Besides, we observed a main effect of music in middle electrodes regions ($F = 16.843$, $p = 0.000$, $\eta^2 = 0.313$). No interactive effects were found between music and nouns (lateral: $F = 0.934$, $p = 0.340$, $\eta^2 = 0.025$; middle: $F = 1.97$, $p = 0.659$, $\eta^2 = 0.005$). To further explore the influence of different music conditions on animate and inanimate nouns, we conducted a simple-effect test, with results showing that inanimate nouns induced greater N400 than the animate nouns in accelerating music conditions (lateral: $F = 13.011$, $p = 0.001$, $\eta^2 = 0.26$) and decelerating music conditions (middle: $F = 5.024$, $p = 0.031$, $\eta^2 = 0.12$). A simple test on the effect of music on nouns further revealed that only action verbs triggered an intensifying N400 in accelerating music than in decelerating music (lateral: $F = 5.574$, $p = 0.024$, $\eta^2 = 0.131$), but inanimate nouns did not show conspicuous processing difference in the two priming-music conditions. In addition, there were significant interactions between noun and region factor ($F = 9.931$, $p = 0.002$, $\eta^2 = 0.212$) and between noun and hemisphere factor ($F = 7.572$, $p = 0.009$, $\eta^2 = 0.17$) in lateral electrodes. No marked three

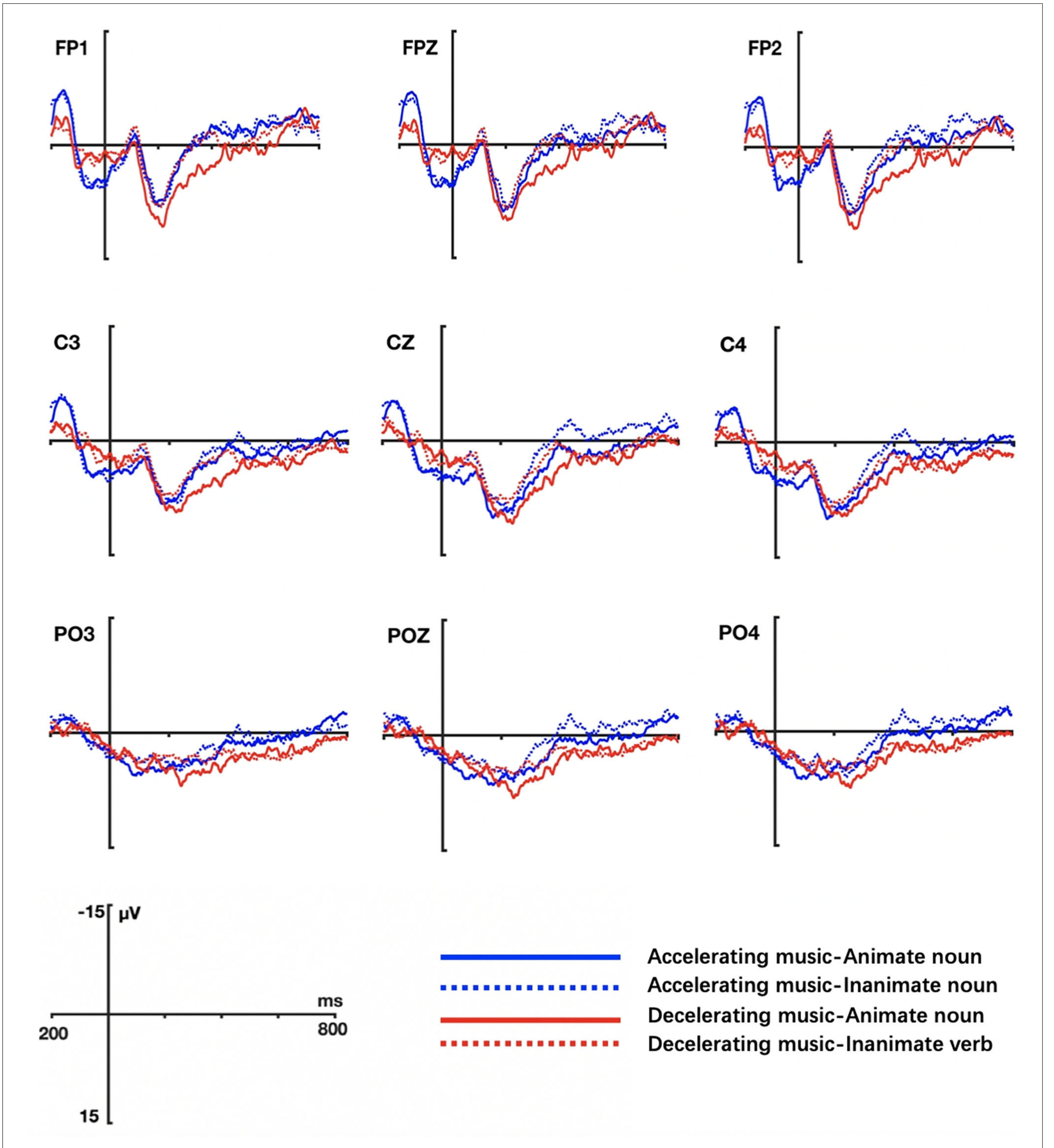

FIGURE 6 | Grand ERP averages elicited by animate and inanimate nouns in tempo-with-change music priming conditions (Task 3).

interaction among music, noun, hemisphere (lateral: $F = 2.656$, $p = 0.112$, $\eta^2 = 0.067$) and music, noun, region (lateral: $F = 0.138$, $p = 0.738$, $\eta^2 = 0.004$; middle: $F = 0.634$, $p = 0.451$, $\eta^2 = 0.017$) were triggered.

In Task 4, there was no main effects in noun type and music type both in lateral (Noun: $F = 0.002$, $p = 0.966$, $\eta^2 = 0.000$; Music: $F = 0.032$, $p = 0.858$, $\eta^2 = 0.001$) and middle electrodes (Noun: $F = 0.054$, $p = 0.818$, $\eta^2 = 0.001$; Music: $F = 0.026$,

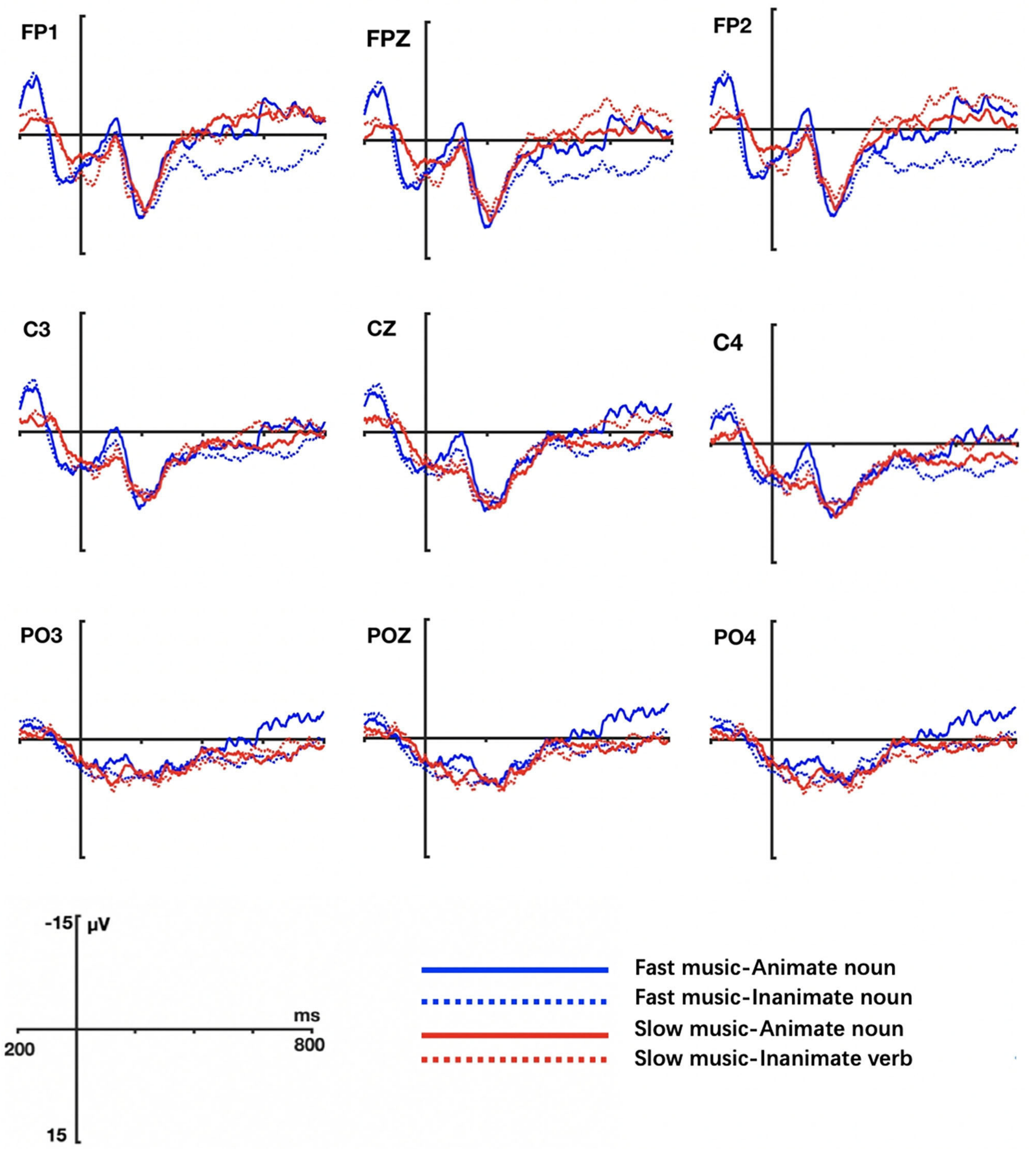

FIGURE 7 | Grand ERP averages elicited by animate and inanimate nouns in tempo-without-change music priming conditions (Task 4).

$p = 0.874$, $\eta^2 = 0.001$). No significant interactions were found between music without tempo changes and nouns (lateral: $F = 0.213$, $p = 0.647$, $\eta^2 = 0.006$; middle: $F = 0.725$, $p = 0.400$, $\eta^2 = 0.019$). These results suggest that the N400 amplitudes induced by the two types of nouns did not differ significantly, nor did those induced by the two types of music. There were no significant interactions between noun and electrode regions factors (lateral: noun × region: $F = 0.158$, $p = 0.763$, $\eta^2 =$

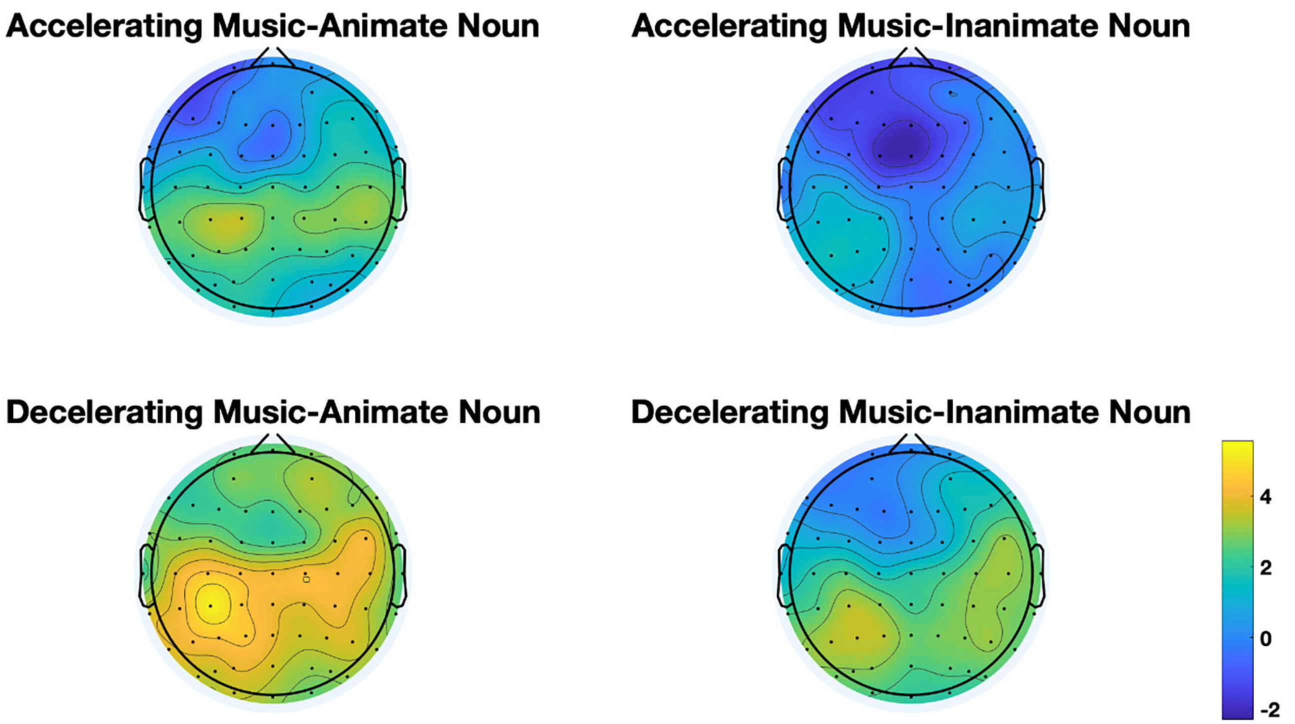

FIGURE 8 | Scalp topographic maps of N400 across the four conditions in Task 3.

0.004; noun × hemisphere: $F = 0.09$, $p = 0.926$, $\eta^2 = 0.000$; middle: noun × region: $F = 0.839$, $p = 0.395$, $\eta^2 = 0.022$) and music and electrode regions factors (lateral: music × region: $F = 2.608$, $p = 0.112$, $\eta^2 = 0.066$; music × hemisphere: $F = 0.892$, $p = 0.351$, $\eta^2 = 0.024$; middle: music × region: $F = 0.713$, $p = 0.426$, $\eta^2 = 0.018$). However, there occurred significant interactions among music, noun, hemisphere (lateral: $F = 10.663$, $p = 0.002$, $\eta^2 = 0.224$) and among music, noun, region (middle: $F = 0.479$, $p = 0.031$, $\eta^2 = 0.112$).

Discussion

This study investigated the processing of nouns primed by different music tempo types. As expected, the results of Tasks 3 and 4 displayed that music with tempo changes influenced the processing of nouns while music without tempo changes did not, which is compatible with Hedger et al.'s (2013) study. This result demonstrated that tempo changes in music were more related to motion concept and triggered the cognition of nouns as a result. Therefore, participants could metaphorically establish the relations of the motion concept between this type of music and nouns on the basis of embodiment. However, this result was not observed in Task 4, suggesting that music without tempo changes can not differentiate between animate nouns and inanimate nouns, as the N400 amplitudes induced by the two types of nouns were not strikingly different. This finding appeared to justify an important recognition that sensorimotor experience was activated in music with tempo changes and nouns in some way, indexed by N400, consistent with our first prediction. Furthermore, this result extends the finding of Task 1, i.e., music with tempo changes influences not only action verbs but also animate nouns. Consequently, we infer that the participants can hardly map the relations between music without tempo changes and nouns (or verbs), as indicated by the absence of different N400 amplitudes and that motion information may not be an integral component of this music type in question, at least off stage in the processing of nouns.

Roughly like the case in Study 1, the processing advantages of congruent pairs failed. To be specific, the increased N400 amplitudes elicited by animate nouns was larger in accelerating music condition than in decelerating music condition, inconsistent with our second prediction. So to speak, the similar effect of music with tempo changes occurred on action verbs in Task 1 and on animate nouns in Task 3. In terms of the word-repetition effect (Michael, 1985) mentioned in Study 1, we assumed that it was the re-occurrence of motion information that led to the N400 amplitudes of animate nouns, for the information was shared by accelerating music and animate nouns.

The electrophysiological results showed that the greater N400 was induced by inanimate nouns (in Task 3), indicating the greater difficulty in processing inanimate nouns vs. animate nouns. Furthermore, in Task 3, the results that music with tempo changes could affect only animate nouns but inanimate nouns, implying that the motion concept correlation between the music with tempo changes and animate nouns is easier to establish. As a result, the sensorimotor experience can be harder to activate during the processing of inanimate nouns, or inanimate nouns are not conceptually related to the motion concept at all, favoring the third prediction.

Above all, Study 2 shows that the embodied experience can be activated in music with tempo changes and animate nouns, indexed by N400. The processing of animate nouns is easier to process than inanimate nouns due to the more motion information they have, favoring our prediction three. There is no motion congruency effect of N400 in music and language, disfavoring the second prediction.

GENERAL DISCUSSION

The two ERP studies aim to investigate how verbs and nouns are processed under two music priming conditions by means of a cross-modality priming paradigm (music listening followed by words' visual reading). Results show that music types influence different classes of words uniquely, action verbs and animate nouns contrast with state verbs and inanimate nouns in processing patterns, and the motion concept as a high-level meaning can be well-established between music tempos and words on the embodied metaphor basis, indexed by N400. The processing of words varies with music tempos, which is not determined by the congruency between music and words but influenced by the properties of music tempos and words.

Tempo-Change Music Affecting Action Verbs and Animates Nouns

The experimental results show that music with tempo changes facilitates the processing of action verbs and animate nouns as revealed by the attenuated ERP amplitudes of the words, consistent with our first prediction. Yet as the priming stimuli are transformed into music without tempo changes, verbs show just the opposite pattern while nouns are not differently affected. The different N400 patterns between the two tasks (Task 1 vs. 2; Task 3 vs. 4) are supposed to result from the attributes of music tempos, suggesting that motion concept is encoded to different degrees in both music and words, i.e., more motion information in music with tempo changes and action verbs and animate nouns while less motion information in music without tempo changes and state verbs and inanimate nouns. According to Hedger et al. (2013), music tempos are closely related to the different properties of speed representations. For accelerating and decelerating music, the number of notes becomes either more or less dense within a single music excerpt, wherein individuals can perceive and expect its speed changes. That is to say, music with tempo changes has strong energy (Eitan and Granot, 2006), hence embracing strong-motion motivation. In statically fast and slow music motif conditions, however, tempos within a music motif have no temporal changes. Here, "fast" and "slow" can be labeled only by comparing one music excerpt with the other. Consequently, compared with statically fast and slow music, the properties of accelerating and decelerating music give people certain clues to identify dynamic changes and anticipate motion changes.

In parallel to the music with tempo changes, action verbs and animate nouns are encoded with more motion information than state verbs and inanimate nouns, giving rise to the reduced N400 effect in words as a result. By contrast, the music without tempo changes is associated with the state rather than motion in concept, and its lack of motion fails to trigger the motion concept in words, resulting in no ERP amplitudes in verbs and nouns as well as their subcategories. That may also explain why people cannot well-establish the relations between music without tempo changes and words (verbs and nouns), to be elaborated below.

Motion Concept-Based Congruency Between Music and Words

In the experiments, participants' responses were not influenced by compatibility between music with or without tempo changes and words (verbs and nouns). The unexpected results that congruent pairs elicited enhanced N400 in Tasks 1, 2, and 3 but both congruent and incongruent pairs induced no amplitude difference in Task 4 to suggest that accelerating music and fast music inhibit the processing of action verbs and animate nouns, seemingly contradictory to Zhou et al.'s (2015) study that incongruency between music and pictures enhanced larger N400. The following makes an attempt to provide several possible explanations for the issue.

The first is due to the greater difficulty of the motion concept in verbs and nouns than in non-verbal stimuli (e.g., pictures). In Hedger et al.'s (2013) study, accelerating music and decelerating music better-primed pictures in motion and at rest, respectively. Compared with pictures, words seem more difficult to have the motion concept extracted, for the decoding of the motion concept in words is involved a more complex process. In general, visual pictures involve a visual perception system at a low level (Kamio and Toichi, 2000; Ostarek and Hüettig, 2017), whereas language processing is an advanced activity of human cognition. According to the spreading activation modal (Collins and Quillian, 1970) that each word related to an entry in the mental lexicon may be activated as one hears a word. We assume that decoding the motion concept of words, one of the lexical aspects of the mental lexicon, may be influenced by other "competitors" like literal meaning and figurative meaning. Compared with language, an image depicted in the scene is more salient and thus restricts other alternatives and "competitors" that may be activated in the mental lexicon (Altmann and Kamide, 1999) during language comprehension. By contrast, the motion concept in the language is not directly encoded and becomes slowly extracted or decoded during linguistic comprehension. As a result, establishing the congruency seems more difficult between music and words than between music and pictures.

The second explanation is related to the word repetition effect. It is discovered that repeated words display more negative-going ERP amplitudes than non-repeated but related words (Michael, 1985). In our experiments, both words (action verbs and animate nouns) and music (accelerating music and fast music) have high motion energy and attributes. Probably, it is the repeated motion information by action verbs and animate nouns that evoked the larger N400. If it is so, the repetition effect is not confined to concrete words with the same form but can be extended to the abstract notion like the motion concept implied in words.

The third reason may be due to the relatively small number of materials adopted in our experiments. Although our studies adopted the same number of musical and word materials as Hedger et al. (2013), the trials are actually less than those in many purely linguistic ERP experiments, which may partly explain the discrepancy between linguistic priming paradigm (within the same domain) and music-linguistic priming paradigm (across domains).

Processing Patterns Differentiating Between Word Types

The two studies flag the different processing patterns of the four types of words (action verbs, state verbs, animate nouns, and inanimate nouns). As tempo-change music serves as the prime, action verbs and animate nouns show easier processing patterns than state verbs and inanimate nouns, respectively, both in accelerating music and decelerating music condition, adding evidence to the view that motion concept is metaphorically established across categories *via* our bodily experience. From the words' encoding perspective, action verbs and animate nouns are associated with more motion information while state verbs and inanimate nouns are often deemed as words with low motion (Muraki et al., 2020). State verbs and inanimate nouns induce larger N400 amplitudes when primed by music with tempo changes but do not interact with the music sub-types, suggesting that they convey less motion information than action verbs and animate nouns. In light of embodiment theories, semantic processing is wholly reliant on sensorimotor experience (Lakoff and Johnson, 1999, p. 151; Glenberg and Kaschak, 2002), our bodily experience non-technically. Likewise, the comprehension of the motion concept in words also cannot be separated from embodied experience. According to Grossman et al. (2002), there is some degree of sensorimotor representation during motion-verb processing while the sensorimotor systems are not involved in the abstract verb representation. Therefore, compared with state verbs, action verbs are more prominent in motion information, which may yield the easier processing of action verbs relative to state verbs in our studies. That is, verbs with more motion information (e.g., action verbs) can be better understood with the help of our previous bodily experience. The processing difficulty of inanimate nouns in our motion-related task extends the previous studies that inanimate nouns in subject position induced larger N400 amplitudes as they are not the ideal actor to perform action relative to animate nouns (Dahl, 2008; Philipp et al., 2008). This finding is consistent with Philipp et al.'s (2008) claim that animate nouns as the typical agent have the ability to perform action whereas inanimate nouns often serve as the undergoer and with Muraki et al.'s (2020) finding that abstract non-bodily state verbs elicited greater N400 amplitudes (at frontocentral electrodes) relative to abstract mental state and concrete verbs in a syntactic classification task.

In no-tempo-change music conditions, only verbs show their differences in their subcategories while the processing of nouns does not vary with fast and slow music conditions, suggesting verbs and nouns are different distinguished by this kind of music. Cognitively, verbs and nouns represent dynamic and static characteristics in image schema, respectively (Li, 2007), taking on different neural basis. From the perspective of processing, verbs are related to the motor cortex and the frontal lobe (near the motor cortex), and nouns are associated with the visual cortex and temporal occipital area (near the visual center) (Pulvermüller et al., 1999). That is to say, verbs involve more motion information than nouns in processing. In our studies, verbs are more sensitive to music without tempo changes than nouns probably owing to the sensorimotor experience that has been activated. A similar difference also finds expressions in the loci involved in processing the two classes of words, i.e., nouns are mainly distributed in the anterior and central regions whereas verbs only activate the anterior regions, indicating verbs are more frontally-distributed in processing. In their sub-categories, action verbs and animate nouns induced larger amplitudes both in fast and slow music conditions suggesting that music without tempo changes inhibited these kinds of verbs. Besides, the N400 amplitudes of animate nouns and inanimate nouns are not affected by music without tempo changes. These reversed results demonstrate that music with tempo changes fail to facilitate the processing of verbs and nouns, probably owing to the different role of music tempo types, as illustrated in section Motion Concept-Based Congruency Between Music and Words.

N400 Indexing Motion Concept by Music and Words

In general, N400 is an electrophysiological index responsible for semantic violation in both languages (e.g., Kutas and Hillyard, 1980; Kutas and Federmeier, 2011) and music (e.g., Koelsch et al., 2004). Moreover, the N400 effects basically keep constant in language conditions (in which target words are preceded by a sequence of words or sentences) and in music conditions (in which target words are preceded by musical excerpts) (Koelsch, 2011). Music is different from words in meaning representation, but they share the same conceptual meaning—the motion concept, which is reliant on sensorimotor experience. In the present study, the motion concept as the extra-musical meaning in music was perceived auditorily but the metaphorical meaning in words was understood visually. The result that action verbs and animate nouns can well-represent the motion concept under tempo-change music conditions demonstrates that the motion concept is semantic in nature, and the resulting N400 is believed to be the electrophysiological indicator of the shared motion by music and words. Our studies are supportive of the hypothesis that music and language share the neural mechanism (N400) of meaning processing.

According to the perceptual symbol's hypothesis (Barsalou, 1999), the storage of perceptual symbols in memory based on perception is extracted from experience (e.g., basic perception like color, emotion, figure, orientation, and motion) and then a simulator will be established. During simulation, a basic functional conceptual system can be constructed, representing categorization, propositions, abstract concepts, and their kind. Therefore, we assume that the perceptual symbols that involve the perceptive motion in music and words can be activated just because our corresponding sensorimotor experience comes into play in understanding the motion concept implicated in words. On this account, action verbs and animate nouns can more easily trigger our sensorimotor experience in the brain so as to make the processing at ease in the motion relatedness task. Music, a motion-based art that unfolds over time, surely contains certain motion information. Hence, we may infer that action verbs and animate nouns (relative to state verbs and inanimate nouns) are easier to process in the music priming condition just on account of the more shared motion information (revealed by the semantic index N400), for sensory-motor cortex may have already been activated as soon as the music is heard by participants. Our results, along with other related studies (Stanfield and Zwaan, 2001; Zwaan et al., 2002; Wang and Zhao, 2020), support the perceptual symbol theory that perceptual symbols of referent are activated during language comprehension (Barsalou, 1999).

Motion Concept Modulated by Embodiment

In our experiments, the motion concept is taken as the referencing point to examine how different music may exert impacts on verbs and nouns. The experimental design in this way is based on the belief that music and words are encoded with the shared motion concept (though in a different manner), a high level of meaning across modalities. The results justified the belief, as indicated by the N400 effect. We argue that the different processing patterns associated with the four classes of words and music are attributed to the diverse degree of embodiment involved in the two domains (music and words).

As demonstrated above, the music with tempo changes is connected with more musical motion, and the music without tempo change with less musical motion whereas action verbs and animate nouns are judged with more motion information than state verbs and inanimate nouns. As a consequence, a decreased N400 was observed in action verbs and animate nouns under the priming condition of music with tempo changes, but verbs showed reversed pattern and nouns showed no obvious N400 difference under the priming condition of music without tempo changes. In our everyday life, physical motion is easily and frequently embodied and encoded as a motion concept in both music and words metaphorically. In association with our results, it appears safe to infer that the more embodied motion is involved in words, the smaller the evoked N400 amplitudes are. Such an inference is favored by few studies indirectly, e.g., the motion concept in music will create an expectation of the target pictures even though it is irrelevant to the task (Hedger et al., 2013; Zhou et al., 2015). The motion concept as a high-level meaning exists across domains, indexed by the N400 effect electrophysiologically, which is in accordance with Zhou et al.'s (2015) study. Just like the case in pictures, the motion concept of verbs and nouns can be well-comprehended metaphorically by virtue of individuals' sensorimotor experience.

In light of embodiment theories, semantic processing is wholly reliant on sensorimotor experience (Lakoff and Johnson, 1999, p. 151; Glenberg and Kaschak, 2002), our bodily experience non-technically. Likewise, the comprehension of the motion concept in words also cannot be separated from embodied experience. Given that our participants performed a motion relatedness task, processing words in our studies was not decoding the literal meaning but interpreting the motion clues of words. Our results are consistent with former studies in which perceptive information (such as orientation, motion features, location, and shape of objects) encoded by language was understood by embodied experience (e.g., Stanfield and Zwaan, 2001; Zwaan et al., 2002; Wang and Zhao, 2020).

Overall, our results indicate that the motion concept is shared by music and words on the embodiment basis and this sharing is indexed by the N400 component, adding further evidence to the view that music and language are supposed to share the neural mechanism of meaning processing (Koelsch et al., 2004; Koelsch, 2011; Jiang, 2016).

CONCLUSION

Music and language signal two domains human beings use to exchange intentions and meanings despite their key differences (specificity, compositionality, and communicativeness) (Slevc and Patel, 2011). The current study demonstrates that music and language are conceptually or semantically correlated with the general motion concept *via* embodiment. Our results obtained indicate that the processing of verbs and nouns is influenced by music tempo types and N400 indexes the motion concept shared by music and words. On the one hand, action verbs and animate nouns are easier to process than state verbs and inanimate nouns, respectively, due to the more motion information they own in the motion-related tasks. On the other hand, four types of words can be distinguished perceptually when primed by different music tempos. Moreover, these findings demonstrate that the motion concept exists regardless of the modalities and domains the priming stimuli and target stimuli are involved in, favoring the embodiment theory in the comprehension of musical and linguistic meaning.

In conclusion, our experiments extend the previous studies by showing that music tempos can conceptually prime words in the language and adds evidence to the hypothesis that music and language share the neural mechanism of meaning processing. The challenge for future research is to explore the neural mechanism of other aspects of meaning assumed to be shared by music and language.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Biomedical Ethics Committee of Qufu Normal University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TZ, YL, and HL conceived the study. YL and TW performed the experiments. YL collated and analyzed the data. YL drafted the first manuscript, which TZ and SZ revised. All authors edited the final version of the manuscript and have approved it for publication.

FUNDING

This work was supported by the Fundamental Research Funds for the Central Universities under the No. 2242022R10100.

ACKNOWLEDGMENTS

Warmest thanks to all the teachers and students in the School of Translation Studies at the University. Also, genuine thanks should be given to all the participants who took part in the ERP experiments and those who helped us conduct a series of questionnaires.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.888226/full#supplementary-material

REFERENCES

Altmann, G. T. M., and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition.* 73, 247–264. doi: 10.1016/S0010-0277(99)00059-1

Amsel, B. D., and Cree, G. S. (2013). Semantic richness, concreteness, and object domain: an electrophysiological study. *Can. J. Exp. Psychol.* 67, 117–129. doi: 10.1037/a0029807

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/S0140525X99002149

Chen, C.-L. (2002). *The Syntactic and Semantic Study on Modal Verbs in Chinese*. Shanghai: Xuelin Press.

Chiang, J. N., Rosenberg, M. H., Buffford, C. A., Stephens, D., Lysy, A., and Monti, M. M. (2018). The language of music: common neural codes for structured sequences in music and natural language. *Brain Lang.* 185, 30–37. doi: 10.1016/j.bandl.2018.07.003

Collins, A. M., and Quillian, M. R. (1970). Facilitating retrieval from semantic memory: the effect of repeating part of an inference. *Acta Psychol.* 33, 304–314. doi: 10.1016/0001-6918(70)90142-3

Dahl, Ö. (2008). Animacy and egophoricity: grammar, ontology and phylogeny. *Lingua* 118, 0–150. doi: 10.1016/j.lingua.2007.02.008

Dai, Y.- J., Hu, Y.-S., and Fan, X. (1995). *Study on Verbs.* Luoyang: Henan University Press.

Daltrozzo, J., and Schön, D. (2009a). Conceptual processing in music as revealed by N400 effects on words and musical targets. *J. Cogn. Neurosci.* 21, 1882–1892. doi: 10.1162/jocn.2009.21113

Daltrozzo, J., and Schön, D. (2009b). Is conceptual processing in music automatic? An electrophysiological approach. *Brain Res.* 1270, 88–94. doi: 10.1016/j.brainres.2009.03.019

de Vega, M., de, Glenberg, A., and Graesser, A. (2008). *Symbols and Embodiment: Debates on Meaning and Cognition*. New York, NY: Oxford University Press.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods.* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Eitan, Z., and Granot, R. Y. (2006). How music moves: musical parameters and listeners images of motion. *Music Percept.* 23, 221–247. doi: 10.1525/mp.2006.23.3.221

Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychonom. Bull. Rev.* 9, 558–565. doi: 10.3758/BF03196313

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., et al. (2002). Neural representation of verb meaning: an fMRI study. *Hum. Brain Mapp.* 15, 124–134. doi: 10.1002/hbm.10117

Hauk, O., Johnsrude, I., and Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9

Hedger, S. C., Nusbaum, H. C., and Hoeckner, B. (2013). Conveying movement in music and prosody. *PLoS ONE* 8, e76744. doi: 10.1371/journal.pone.0076744

Horchak, O. V., Giger, J. C., Cabral, M., and Pochwatko, G. (2014). From demonstration to theory in embodied language comprehension: a review. *Cogn. Syst. Res.* 29–30, 66–85. doi: 10.1016/j.cogsys.2013.09.002

Hu, Z. -L., Zhu, Y. -S., and Zhang, D. -L. (1989). *A Survey of Systemic-Functional Grammar*. Changsha: Hunan Education Press.

Jentschke, S., Friederici, A. D., and Koelsch, S. (2014). Neural correlates of music-syntactic processing in two-year old children. *Dev. Cogn. Neurosci.* 9, 200–208. doi: 10.1016/j.dcn.2014.04.005

Jiang, C.-M. (2016). *Psychology of Music*. Shanghai: East China Normal University Press.

Johnson, M. L., and Larson, S. (2003). "Something in the way she moves": metaphors of musical motion. *Metaphor Symbol* 18, 63–84. doi: 10.1207/S15327868MS1802_1

Kamio, Y., and Toichi, M. (2000). Dual access to semantics in autism: Is pictorial access superior to verbal access? *J. Child Psychol. Psychiatry* 41, 859–867. doi: 10.1111/1469-7610.00673

Koelsch, S. (2011). Towards a neural basis of processing musical semantics. *Phys. Life Rev.* 8, 89–105. doi: 10.1016/j.plrev.2011.04.004

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nat. Neurosci.* 7, 302–307. doi: 10.1038/nn1197

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 27, 4427. doi: 10.1126/science.7350657

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Larson, S. (2002). Musical forces, melodic expectation, and jazz melody. *Music Percept.* 19, 351–385. doi: 10.1525/mp.2002.19.3.351

Larson, S. (2004). Musical forces and melodic expectations: comparing computer models and experimental results. *Music Percept.* 21, 457–498. doi: 10.1525/mp.2004.21.4.457

Li, F.-Y. (2007). On image schema theory. *J. Sichuan Int. Stud. Univ.* 23, 6. doi: 10.3969/j.issn.1674-6414.2007.01.017

Ma, Q.-Z. (2005). *The Chinese Verb and Verbal Constructions.* Beijing: Peking University Press.

Michael, D. R. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology* 39, 123–148.

Muraki, E. J., Cortese, F., Protzner, A. B., and Pexman, P. M. (2020). Heterogeneity in abstract verbs: an ERP study. *Brain Lang.* 211, 104863. doi: 10.1016/j.bandl.2020.104863

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Olejnik, S., and Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol. Methods* 8, 434–447. doi: 10.1037/1082-989X.8.4.434

Ostarek, M., and Hüettig, F. (2017). Spoken words can make the invisible visible – Testing the involvement of low-level visual representations in spoken word processing. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 499–508. doi: 10.1037/xhp0000313

Patel, A. D. (2008). *Music, Language, and the Brain.* New York, NY: Oxford University Press.

Pauen, S., and Träuble, B. (2009). How 7-month-olds interpret ambiguous motion events: category-specific reasoning in infancy. *Cogn. Psychol.* 59, 275–295. doi: 10.1016/j.cogpsych.2009.06.001

Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., and Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: evidence from auditory event-related brain potentials. *Brain Lang.* 105, 112–133. doi: 10.1016/j.bandl.2007.09.005

Pulvermüller, F., Lutzenberger, W., and Preissl, H. (1999). Nouns and verbs in the intact brain: evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex* 9, 498–508. doi: 10.1093/cercor/9.5.497

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna.

Savaki, H. E., and Raos, V. (2019). Action perception and motor imagery: mental practice of action. *Prog. Neurobiol.* 175, 107–125. doi: 10.1016/j.pneurobio.2019.01.007

Shao, J.-M., and Liu, Y. (2001). On nouns' dynamicity and the testing approaches. *Chin. Lang. Learn.* 6, 2–7. doi: 10.3969/j.issn.1003-7365.2001.06.001

Slevc, L. R., and Patel, A. D. (2011). Meaning in music and language: three key differences. comment on "Towards a neural basis of processing musical semantics" by Stefan Koelsch. *Phys. Life Rev.* 8,110–111. doi: 10.1016/j.plrev.2011.05.003

Sloboda, J. (1998). Does music mean anything? *Musicae Sci.* 2, 19–31. doi: 10.1177/102986499800200102

Stanfield, R. A., and Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* 12, 153–156. doi: 10.1111/1467-9280.00326

Steinbeis, N., and Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and FMRI provides evidence for similar and distinct neural representations. *PLoS ONE* 3:5434. doi: 10.1371/annotation/a2cfd2c0-5084-4426-8868-f55ec0ea5434

Todd, N. P. M. (1999). Motion in music: a neurobiological perspective. *Music Percept.* 17, 115–126. doi: 10.2307/40285814

Träuble, B., Pauen, S., and Poulin-Dubois, D. (2014). Speed and direction changes induce the perception of animacy in 7-month-old infants. *Front. Psychol.* 5, 1141. doi: 10.3389/fpsyg.2014.01141

Wang, H. L., and Zhao, Y. H. (2020). Embodiment effect of implicit shape information in L2 sentence comprehension. *J. Tianjin For. Stud. Univ.* 27, 70–79. doi: 10.3969/j.issn.1008-665X.2020.05.007

Weckerly, J., and Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology* 36, 559–570. doi: 10.1111/1469-8986.3650559

Wolter, S., Dudschig, C., Irmgard, D., and Kaup, B. (2015). Musical metaphors: evidence for a spatial grounding of non-literal sentences describing auditory events. *Acta Psychol.* 156, 126–135. doi: 10.1016/j.actpsy.2014.09.006

Yuan, M.-J. (1998). Further notes on the classification of non-volitional verbs. *Stud. Chin. Language.* 4, 24–30.

Zhou, L. S., Jiang, C. M., Wu, Y. Y., and Yang, Y. F. (2015). Conveying the concept of movement in music: an event-related brain potential study. *Neuropsychologia* 77, 128–136. doi: 10.1016/j.neuropsychologia.2015.07.029

Zwaan, R. A., Stanfield, R. A., and Yaxley, R. H. (2002). Language comprehenders mentally represent the shape of objects. *Psychol. Sci.* 13, 168–171. doi: 10.1111/1467-9280.00430

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 20 June 2022
doi: 10.3389/fpsyg.2022.894114

Bilingual Processing Mechanisms of Scientific Metaphors and Conventional Metaphors: Evidence *via* a Contrastive Event-Related Potentials Study

Xuemei Tang[1†], Lexian Shen[1†], Peng Yang[1], Yanhong Huang[1], Shaojuan Huang[1], Min Huang[1] and Wei Ren[2*]*

[1] School of Foreign Studies, Anhui Polytechnic University, Wuhu, China, [2] Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an, China

OPEN ACCESS

Edited by:
Huili Wang,
Dalian University of Technology, China

Reviewed by:
Pléh Csaba,
Central European University, Hungary
Xiaodong Xu,
Nanjing Normal University, China

***Correspondence:**
Min Huang
495531005@qq.com
Wei Ren
renwei@snnu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: 11 March 2022
Accepted: 23 May 2022
Published: 20 June 2022

Citation:
Tang X, Shen L, Yang P, Huang Y, Huang S, Huang M and Ren W (2022) Bilingual Processing Mechanisms of Scientific Metaphors and Conventional Metaphors: Evidence via a Contrastive Event-Related Potentials Study. Front. Psychol. 13:894114. doi: 10.3389/fpsyg.2022.894114

To study the different mechanisms of understanding figurative language in a speaker's native language (L1) and their second language (L2), this study investigated how scientific metaphors in Chinese (L1) and English (L2) are electrophysiologically processed *via* event-related potential experimentation. Compared with the metaphors from daily life or in literary works, scientific metaphors tend to involve both a more complicated context structure and a distinct knowledge-inferencing process. During the N400 time window (300–500 ms), English scientific metaphors elicited more negative N400s than Chinese ones at the parietal region. In the late positive component (LPC) time window (550–800 ms), English scientific metaphors elicited less positive LPCs than Chinese ones at the parietal region, and larger late negativities encompassing smaller areas of the brain. The findings might indicate that for late unbalanced bilingual speakers, L2 scientific metaphor comprehension requires more effort in information retrieval or access to the non-literal route. Altogether, the possible findings are that non-native and non-dominant language processing involves decreased automaticity of cognitive mechanisms, and decreased sensitivity to the levels of conventionality of metaphoric meanings.

Keywords: scientific metaphors, bilingual, event-related potentials (ERPs), N400, late negativity

INTRODUCTION

The semantic integration between the daily source domain and the scientific target domain involved in scientific metaphor processing provides a perspective for exploring the similarities and differences between processing mechanisms of native language and non-native language (Jankowiak et al., 2017). However, previous studies focused on conventional metaphors and devoted little attention to scientific metaphors. Such metaphors are frequently deployed by scientists who use concrete and intuitive objects and knowledge to describe abstract concepts or objects that are impossible to observe directly. The mechanism of scientific metaphors reflects the core idea of conceptual metaphor theory, the metaphorical understanding of abstract concepts through more specific knowledge structures (Lakoff and Johnson, 1980). The continuous innovation of natural science requires the reorganization and construction of thinking and expression methods in systems related to new concept expression (Tang et al., 2017a,b). Compared with conventional metaphors, the processing of scientific metaphors involves a unique mechanism involving a complicated contextual structure and knowledge inference, which might provide insight in observing the between-language differences of the brain correlates modulated by nativeness, thus deepening the comparative study of bilingual processing mechanisms. Moreover, the comparative study of two different types of metaphors (scientific metaphors and conventional metaphors) might shed new light on the discussion of the processing mechanism of figurative language.

The processing of scientific metaphors differs from that of conventional metaphors mainly in two aspects. First, the contextual structure of scientific metaphors is more complex than that of conventional metaphors, which has been established in previous N400 experiments (Tang et al., 2017a). For example, the source domain (*code*) of the scientific metaphor (*A mitochondrion is a code*) is more frequently used in everyday contexts, while the target domain (*mitochondria*) is more frequently used in scientific contexts. Second, the late processing period of scientific metaphors probably involves a reasoning process from the daily concrete concept to the scientific abstract concept, so as to better comprehend the embedded scientific knowledge. The late components of ERP [the late positive component (LPC) and the late negativity] are very sensitive to this process.

Event-related potentials (ERPs) with extremely high time resolution and rich component dimensions are increasingly used in the study of metaphorical cognitive neural mechanisms. The N400 is considered to be very sensitive to semantic violations (Kutas and Hillyard, 1980), and is an important index to measure the difficulty in retrieving information stored in the mind. Several monolingual studies on figurative meaning processing reported a modulation of N400 amplitudes by the degree of metaphor conventionality (Pynte et al., 1996; Tartter et al., 2002; Coulson and Van Petten, 2007; Lai et al., 2009). Such results seem to be in line with the Career of Metaphor Model (Bowdle and Gentner, 2005), which postulates that novel metaphors are understood through comparison, while conventional metaphors are preferentially analyzed by categorization. Namely, the processing of novel metaphors involves structure alignment between a metaphoric target and a literal base. In contrast, when processing conventional metaphors, the metaphoric target concept is understood as a member of a superordinate category specified by the literal base term (Jankowiak et al., 2017). Thus, understanding novel metaphors requires engagement in sense creation, while understanding conventional metaphors involves easier sense retrieval, as revealed by enhanced N400 amplitudes for novel metaphors (Arzouan et al., 2007a; Lai et al., 2009). In addition to novelty, N400 is also sensitive to concreteness. Forgács et al. (2014, 2015) proposed that most metaphors describe abstract concepts, which, in their study, elicited larger N400s compared with abstract literal expressions when stimulus novelty was controlled across types. In this study, the target words of scientific metaphors were scientific terms and were relatively more abstract than the daily targets of conventional metaphors, which might contribute to the enhanced N400 amplitudes.

The late positive component (LPC) is considered to reflect integration or reprocessing at the sentence level (Kaan et al., 2000). Some monolingual studies found that novel metaphors elicited greater LPCs than conventional metaphors (De Grauwe et al., 2010; Weiland et al., 2014), which indicates that the semantic processes of metaphorical sentences require additional retrieval of information from semantic memory. Other monolingual studies found that novel metaphors elicited smaller LPCs than conventional metaphors (Arzouan et al., 2007a,b; Zhao et al., 2011; Ma et al., 2016), which might be caused by a late negativity overlapping in the LPC time window, reflecting the secondary processing of further semantic integration. Such late negativity is also considered as a continuation of the N400, indicating a sustained difficulty in fusing two concepts from the source domain and the target domain (Goldstein et al., 2012; Rutter et al., 2012).

So far, few studies focused on the processing mechanism of scientific metaphors for bilingual speakers. Thus, expanding a monolingual electrophysiological study on scientific metaphor processing to include a bilingual component might be insightful to enunciate conceptual integration mechanisms when processing semantically complex meanings. Some bilingual ERP studies have reported a lower N400 amplitude elicited by L2 relative to L1 stimuli (Proverbio et al., 2002; Moreno et al., 2008; Midgley et al., 2009; Newman et al., 2012; Heidlmayr et al., 2015), probably indicating weaker interconnectivity for L2 words within the semantic network compared with L1 words (Midgley et al., 2009). Consequently, when processing non-native language, the activity in long-term memory decreases due to weaker interconnectivity. The functional role of the N400 effect is linked to memory operations involved in information retrieval (Kutas and Federmeier, 2000; Kotz et al., 2012). The current study seeks to build on existing research as to semantic processing in bilinguals by using materials with different semantic complexity ranging from highly complex (scientific metaphoric) and relatively complex (conventional metaphoric) to relatively simple (literal) utterances. To minimize the potential influence of any between-group individual differences, a within-subject design was adopted. Accordingly, contrastive analyses of cognitive mechanisms were carried out to investigate native and non-native figurative language comprehension, which can further reveal how semantic complexity and language dominance interact with each other when processing bilingual languages.

The main objective of this study was to observe brain responses to scientific metaphors, conventional metaphors, and literal expressions in Chinese (L1) and English (L2). First, we aimed to observe whether scientific metaphors would evoke higher N400 amplitudes compared with conventional metaphors in both L1 and L2. That is to say, our first focus was the modulation of conventionality for the N400 effect in two languages. A linear N400 effect has been reported by two previous monolingual studies (Tang et al., 2017a,b), with more pronounced N400 amplitudes for scientific metaphors relative to conventional metaphors. This increased processing difficulty was evident due to more demanding mappings required for scientific metaphor comprehension when complicated context and scientific reasoning are engaged. Second, we aimed to observe whether English scientific metaphors would elicit attenuated N400 effects in comparison to Chinese ones, as has been found in previous bilingual research on semantic processing (Weber-Fox and Neville, 1996; Proverbio et al., 2002; Phillips et al., 2004; Moreno and Kutas, 2005; Midgley et al., 2009; Braunstein et al., 2012; Newman et al., 2012; Heidlmayr et al., 2015). Lower N400s elicited by English materials compared with Chinese ones would further imply weaker semantic interconnectivity for L2 when compared with L1 stimuli (Midgley et al., 2009). Third, we aimed to examine whether language nativeness would modulate cognitive mechanisms involved in semantic reintegration at the later stage of processing, as indexed by the LPC response. Several monolingual studies have reported higher LPC amplitudes as the index of secondary semantic integration (van Herten et al., 2005; De Grauwe et al., 2010; Brouwer et al., 2012; Rataj, 2014). It was expected that English scientific metaphors would elicit decreased LPC amplitudes in comparison to Chinese ones. Lower LPCs elicited by English materials relative to Chinese ones would point to weaker semantic processing for L2 compared with L1 sentences (Newman et al., 2012). Furthermore, if the LPC effect is modulated by the conventionality of metaphors (Arzouan et al., 2007a), scientific metaphors should evoke more pronounced LPC amplitudes than conventional metaphors, indicating the continuation of information retrieval or access to the non-literal route. Finally, we expected that main effects between LPCs across items could be observed in both L1 and L2, indicating a similar sensitivity to different levels of conventionality of metaphors.

EXPERIMENT 1 (CHINESE L1)

Method

Participants

In total, 20 participants (right-handed, healthy, undergraduate students, L1-Chinese) took part in the ERP experiment. They had either normal or corrected vision and no history of mental illness, neurological disorders, or severe brain damage. All subjects started to learn English from elementary school and have passed the CET-6. All subjects signed a consent and confidentiality agreement before the experiment and received remuneration after the experiment. Finally, the trial data of three subjects were eliminated due to failure to meet the 80% completion-rate threshold. Therefore, the final number of subjects included in our statistical analysis was 17 (7 men, 10 women, average age 21.3 ± 5.32 years).

Stimuli

The stimulus pool consisted of 120 sentences, which fell into three categories, namely, scientific metaphors, conventional metaphors, and literal sentences, with 40 sentences in each sentence category. This pool matched the one used in our previous study (Tang et al., 2017a,b) (refer to **Table 1** for details).

TABLE 1 | Sample stimuli.

Scientific metaphors	电路/是/阶梯。	dianlu/shi/jieti.
	离子/是/碎片。	lizi/shi/suipian.
	函数/是/斜坡。	hanshu/shi/xiepo.
	螺栓/是/鱼尾。	luoshuan/shi/yuwei.
	声音/是/波浪。	shengyin/shi/bolang.
Conventional metaphors	恋爱/是/咖啡。	lianai/shi/kafei.
	杭州/是/天堂。	hangzhou/shi/tiantang.
	家庭/是/港湾。	jiating/shi/gangwan.
	手机/是/伙伴。	shouji/shi/huoban.
	语言/是/桥梁。	yuyan/shi/qiaoliang.
Literal expressions	教授/是/学者。	jiaoshou/shi/xuezhe.
	汉语/是/语言。	hanyu/shi/yuyan.
	北京/是/首都。	beijing/shi/shoudu.
	蚂蚁/是/昆虫。	mayi/shi/kunchong.
	小狗/是/宠物。	xiaogou/shi/chongwu.

Procedure

The experiment took place in a sound-attenuated, electrically shielded room. Participants were seated 80 cm away from the display screen. The sentences were presented in white color on a black background, word by word in a quasi-random order. As illustrated in **Figure 1**, stimuli in each trial were presented in the following time sequence: fixation cross (800 ms), blank (200–500 ms), subject (1,000 ms), verb (600 ms), blank (200–500 ms), object (1,000 ms), and question mark (3,000 ms). At the sight of the question mark, participants were asked to do a semantic judgment by pressing a corresponding key with right and left index fingers. The whole experiment consisted of four blocks interspersed with three rest intervals. To familiarize the subjects with trial procedure and operation, a practice session was done before the formal experiment. To mitigate any difficulty in understanding the scientific terms involved in the stimuli, participants were first asked to read a list covering all the scientific terms used.

EEG Recording and Data Analysis

Scalp voltages were collected using the CURRY 7 system (Compumedics Neuroscan, Texas, United States) with 64 Ag/AgCl electrodes, monitored using the CURRY recording software and connected to a SynAmp amplifier (Compumedics Neuroscan, Texas, United States). Amplified analog voltages were digitized at 1,000 Hz. Impedances of individual sensors were kept below 5 kΩ. Eye movements were monitored through bipolar electrodes, which were placed above and below the right eye, as well as at the left and right canthi. Electroencephalography (EEG) was measured online with reference to the left mastoid, with a ground electrode on the medial frontal aspect, and later was analyzed offline with re-reference to an average of the left and right mastoids.

EEG was analyzed using the SCAN 4.5 software (Compumedics Neuroscan, Texas, United States) and Matlab using the ERPLAB toolbox (Lopez-Calderon and Luck, 2014). The EEG was digitally filtered at 0.1–30 Hz bandpass. Eye movements were corrected with an ocular artifact correction algorithm (Gratton et al., 1983). Artifacts with amplitudes

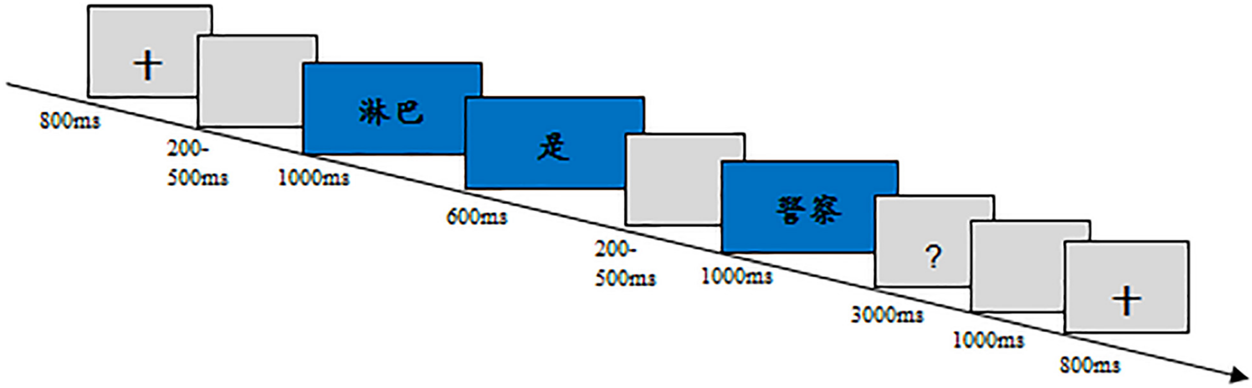

FIGURE 1 | The procedure of Experiment 1.

exceeding ±75 μV were removed from analyses. ERPs were time-locked to the onset of the last word of the sentence and were obtained by stimulus-locked averaging of the EEG recorded in each condition. Epochs were 1,000 ms in length with a 200 ms pre-stimulus baseline. The resulting amplitudes of N400 and LPC were entered into 3 condition × 3 region (frontal F3, Fz, F4, central C3, Cz, C4, parietal P3, Pz, P4) × 3 hemisphere (left F3, C3, P3, midline Fz, Cz, Pz, right F4, C4, P4) three-way ANOVAs for repeated measures. All ANOVA results were Greenhouse–Geisser corrected if assumption of sphericity was violated, and *post-hoc* multiple comparisons were carried out using Bonferroni-adjusted corrections.

Results

Behavioral Results

A repeated-measures ANOVA revealed significant effects of condition for reaction time $F(2,32) = 5.32$, $p = 0.01$, $\eta_p^2 = 0.25$. Pairwise comparisons showed that the reaction time of scientific metaphors was significantly longer than that of literal sentences ($p = 0.005$). There was neither significant difference between conventional metaphors and literal sentences nor between the two metaphorical conditions (p-values > 0.1). For accuracy rates, a main effect between conditions was found [$F(2,32) = 31.12$, $p < 0.001$, $\eta_p^2 = 0.49$]. Pairwise comparisons showed that the accuracy rate of scientific metaphors was significantly lower than that of conventional metaphors and literal sentences (p-values < 0.01). There was no significant difference between conventional metaphors and literal sentences ($p = 0.569$).

Electrophysiological Results

According to the Grand average ERP waveforms recorded at the nine chosen electrodes (see **Figure 2**), there was a sizable negative deflection at about 400 ms, identified as N400, and a later positive deflection appeared from 550 to 800 ms, identified as LPC.

300–500 ms

The condition × region × hemisphere ANOVA revealed a significant main effect of condition [$F(2,32) = 12.56$, $p < 0.001$, $\eta_p^2 = 0.44$]. Scientific metaphors elicited the most negative N400 ($M = 0.61$, $SD = 3.49$), followed by conventional metaphors ($M = 1.75$, $SD = 3.64$) and literal sentences ($M = 2.87$, $SD = 4.63$). Pairwise comparisons revealed a significant difference between N400s elicited by scientific metaphors and conventional metaphors [$t(16) = -2.26$, $p = 0.038$]. Meanwhile, both scientific metaphors and conventional metaphors elicited more negative N400s than literal sentences [scientific metaphors: $t(16) = -6.46$, $p < 0.001$; conventional metaphors: $t(16) = -2.32$, $p = 0.034$].

550–800 ms

Late Positive Component. The main effect of condition was found to be significant [$F(2,32) = 5.22$, $p = 0.024$, $\eta_p^2 = 0.25$]. The lowest LPC amplitude was registered by scientific metaphors ($M = 1.25$, $SD = 3.94$) followed by conventional metaphors ($M = 3.05$, $SD = 4.26$) and literal sentences ($M = 3.44$, $SD = 4.11$). Pairwise comparisons showed that there were significant differences between LPCs elicited by scientific metaphors and conventional metaphors [$t(16) = -2.33$, $p = 0.033$] and by scientific metaphors and literal sentences [$t(16) = -5.27$, $p < 0.001$] but no significant difference between conventional metaphors and literal sentences [$t(16) = -0.44$, $p = 0.67$].

Late Negativity. In addition, the Grand average ERP waveforms of the differences between metaphorical and literal conditions (see **Figure 3**) showed that both scientific metaphors and conventional metaphors elicited a late negativity in the LPC time window. The 2 × 3 × 3 ANOVA revealed a significant main effect of condition [$F(1,16) = 5.41$, $p = 0.033$, $\eta_p^2 = 0.25$]. Scientific metaphors elicited a significantly larger late negativity ($M = -2.19$, $SD = 3.51$) than conventional metaphors ($M = -0.39$, $SD = 4.48$).

Discussion

Consistent with our predictions, scientific metaphors elicited more negative N400 readings than conventional metaphors. The processing of scientific metaphors involves the conceptual integration between scientific-target and daily-source, while the processing of conventional metaphors involves the conceptual integration of daily-target and daily-source, resulting in the contextual complexity found in scientific metaphors. Compared with conventional metaphors, when processing scientific

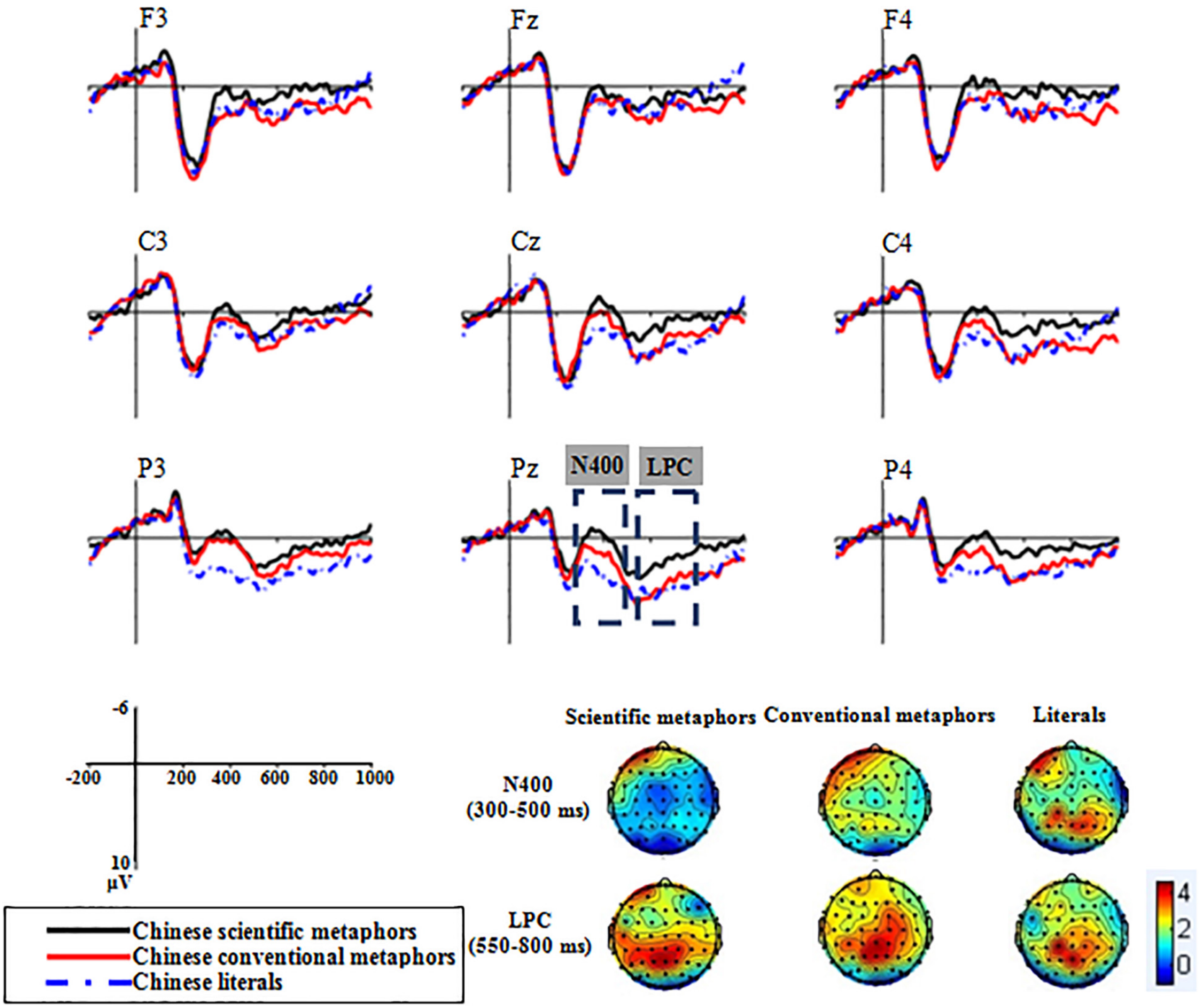

FIGURE 2 | Grand average ERP waveforms of Experiment 1 recorded at the nine chosen electrodes.

metaphors, it might be more difficult to search and retrieve stored conceptual knowledge due to the longer distance between the target and source domains (Tang et al., 2017a,b). Meanwhile, larger N400s elicited by scientific metaphors might also indicate the modulation of concreteness. The scientific targets of scientific metaphors were more abstract than the daily targets of conventional metaphors, contributing to the increased negativities found (Forgács et al., 2015).

Consistent with our predictions, both scientific metaphors and conventional metaphors elicited a late negativity partly overlapping in space and time with the LPC (Arzouan et al., 2007a,b; Zucker and Mudrik, 2019). The higher amplitude of scientific metaphors might be caused by the late inference of scientific metaphors from the daily source domain to the scientific target domain in order to understand the related scientific knowledge (Tang et al., 2017a,b). Compared with conventional metaphors, the late stage of scientific metaphor processing involves deeper secondary semantic integration processes. Moreover, as in novel metaphors, the late processing of scientific metaphors is more taxing on working memory (Steinhauer et al., 2010) and more difficult in semantic integration (Zhao et al., 2011; Goldstein et al., 2012; Rutter et al., 2012).

EXPERIMENT 2 (ENGLISH L2)

Method

Participants

Participants of Experiment 2 were selected from undergraduates with similar ages, language background, and English proficiency as those of Experiment 1. Twenty participants took part in the ERP experiment. As in Experiment 1, the number of subjects included in our final statistical analysis was 17 (8 men, 9 women, average age 21.1 $\pm$ 6.54 years).

Stimuli

The English stimuli of Experiment 2 were the English counterparts of the Chinese stimuli used in Experiment 1

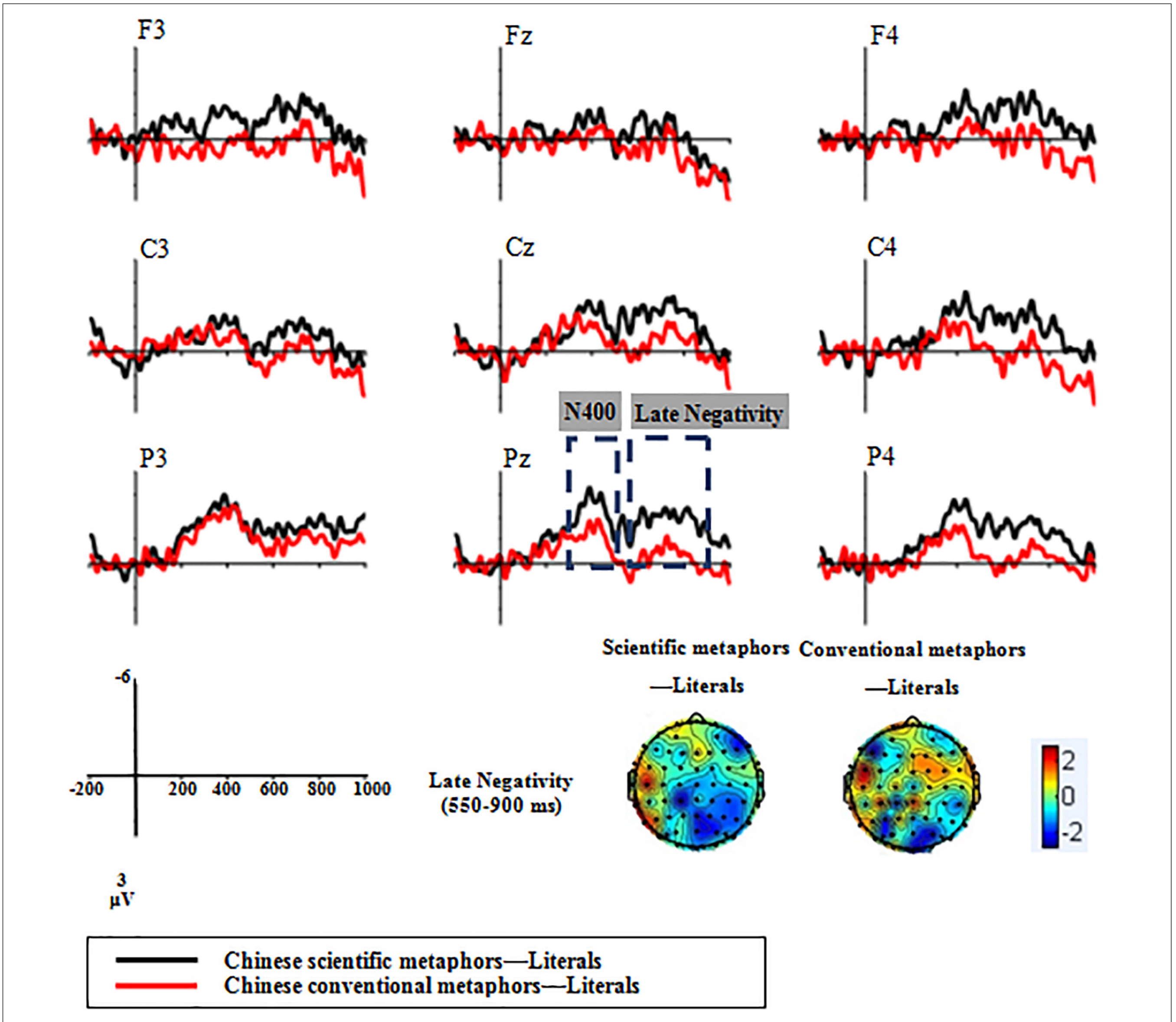

FIGURE 3 | Topographic plots of the two kinds of Chinese metaphors subtracting the literal from the metaphoric in the 300–500 and 550–800 ms time windows.

(see **Table 2**). The English stimuli followed the "A IS B" structure similar to the structure of the Chinese equivalents. As English is a L2 for the subjects, prior to the neurophysiological study, the English stimuli were tested for familiarity by 40 raters who did not participate in the ERP experiment. Before the test, all raters read a pool of English scientific terms used in the scientific sentences. During the test, the raters were asked to decide whether each expression was familiar or not on a 1–5 scale (1 = not familiar, 5 = highly familiar). A total of 120 English stimuli (40 for each category) with a familiarity higher than 3 were chosen for the ERP experiment. A repeated-measures ANOVA revealed a significant main effect of condition [$F(2,78) = 39.72$, $p < 0.001$, $\eta_p^2 = 0.51$]. The familiarity of scientific metaphors was significantly lower than that of conventional metaphors and literal sentences (p-values < 0.001), while no significant difference was found between the familiarity of conventional metaphors and literal sentences ($p = 0.225$).

Procedure

Same as Experiment 1 (see **Figure 4**).

EEG Recording and Data Analysis

Same as Experiment 1.

Results Experiment 2

Behavioral Performance

Diverging from the result of Experiment 1, the main effect of condition for reaction time was not found to be significant ($p = 0.48$). A repeated-measures ANOVA performed on

TABLE 2 | Sample English stimuli.

Scientific metaphors	A charge is flow.
	A conductor is a tunnel.
	A mitochondrion is a code.
	A virus is a killer.
	Sound is wave.
Conventional metaphors	A book is a friend.
	Hangzhou is heaven.
	Language is a bridge.
	Nature is a doctor.
	History is a mirror.
Literal expressions	A professor is a scholar.
	A rose is a plant.
	Beijing is a city.
	Running is a sport.
	Painting is art.

the accuracy rates yielded significant effects of condition [$F(2,32) = 14.73$, $p < 0.001$, $\eta_p^2 = 0.48$]. The difference between the accuracy rates of two metaphorical conditions was not significant ($p = 0.775$), but the accuracy rates of metaphorical conditions were significantly lower than that of the literal condition (p-values < 0.001).

Electrophysiological Results

According to the Grand average ERP waveforms recorded at the nine chosen electrodes (see **Figure 5**), there was a sizable negative deflection at about 400 ms (identified as N400) and a later positive deflection from 550 to 800 ms (identified as LPC).

300–500 ms

As with Chinese metaphor processing, the condition × region × hemisphere ANOVA revealed a significant main effect of condition [$F(2,32) = 20.73$, $p < 0.001$, $\eta_p^2 = 0.56$]. Scientific metaphors elicited more negative N400s ($M = -0.02$, $SD = 3.63$) than both conventional metaphors ($M = 0.79$, $SD = 3.58$) and literal sentences ($M = 2.24$, $SD = 2.99$). Pairwise comparisons showed a significant difference between N400s elicited by scientific metaphors and conventional metaphors [$t(16) = -3.07$, $p = 0.007$]. Meanwhile, both scientific metaphors and conventional metaphors elicited significantly larger N400s than literal sentences [scientific metaphors: $t(16) = -5.07$, $p < 0.001$; conventional metaphors: $t(16) = -4.37$, $p < 0.001$].

In contrast to Chinese metaphor processing, the differences between English scientific metaphors and conventional metaphors were significant at the frontal [$t(16) = -2.88$, $p = 0.011$], central [$t(16) = -2.53$, $p = 0.022$], and parietal [$t(16) = -2.85$, $p = 0.012$] regions. The difference between N400s elicited by English scientific metaphors and conventional metaphors was marginally significant at the left and middle parietal regions [P3: $t(16) = -1.79$, $p = 0.091$; Pz: $t(16) = -1.90$, $p = 0.076$], significant at the right frontal and central regions [F4: $t(16) = -3.33$, $p = 0.004$; C4: $t(16) = -3.50$, $p = 0.003$], and highly significant at the right parietal region [P4: $t(16) = -3.65$, $p = 0.002$]. Relative to English literal sentences, both English scientific metaphors and English conventional metaphors elicited a more negative N400 at all the three regions {scientific metaphor vs. literal sentences: frontal [$t(16) = -4.61$, $p < 0.001$], central [$t(16) = -4.21$, $p = 0.001$], and parietal [$t(16) = -6.13$, $p < 0.001$] regions; conventional metaphors vs. literal sentences: frontal [$t(16) = -3.19$, $p = 0.006$], central [$t(16) = -4.17$, $p = 0.001$], and parietal [$t(16) = -5.09$, $p < 0.001$] regions}.

550–800 ms

Late Positive Component. The condition × region × hemisphere ANOVA revealed a significant main effect of the condition [$F(2,32) = 5.19$, $p = 0.014$, $\eta_p^2 = 0.25$]. Scientific metaphors elicited less positive LPCs ($M = -0.16$, $SD = 3.92$) than conventional metaphors ($M = 1$, $SD = 4.49$) and literal sentences ($M = 1.46$, $SD = 3.42$). The condition × region interaction effect was significant [$F(4,64) = 9.09$, $p < 0.001$, $\eta_p^2 = 0.36$]. Pairwise comparisons showed that there was a marginally significant difference between LPCs elicited by scientific metaphors and conventional metaphors [$t(16) = -1.96$, $p = 0.068$] and a significant difference between scientific metaphors and literal sentences [$t(16) = -3.61$, $p = 0.002$]. Between conventional metaphors and literal sentences, despite an insignificant main effect of condition [$t(16) = -0.91$, $p = 0.37$], the condition × region interaction effect was significant

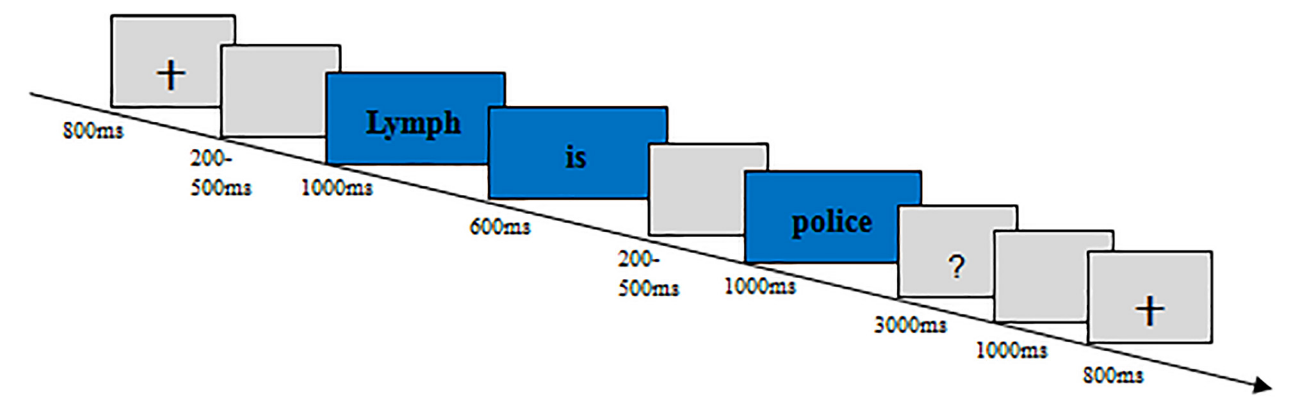

FIGURE 4 | The procedure of Experiment 2.

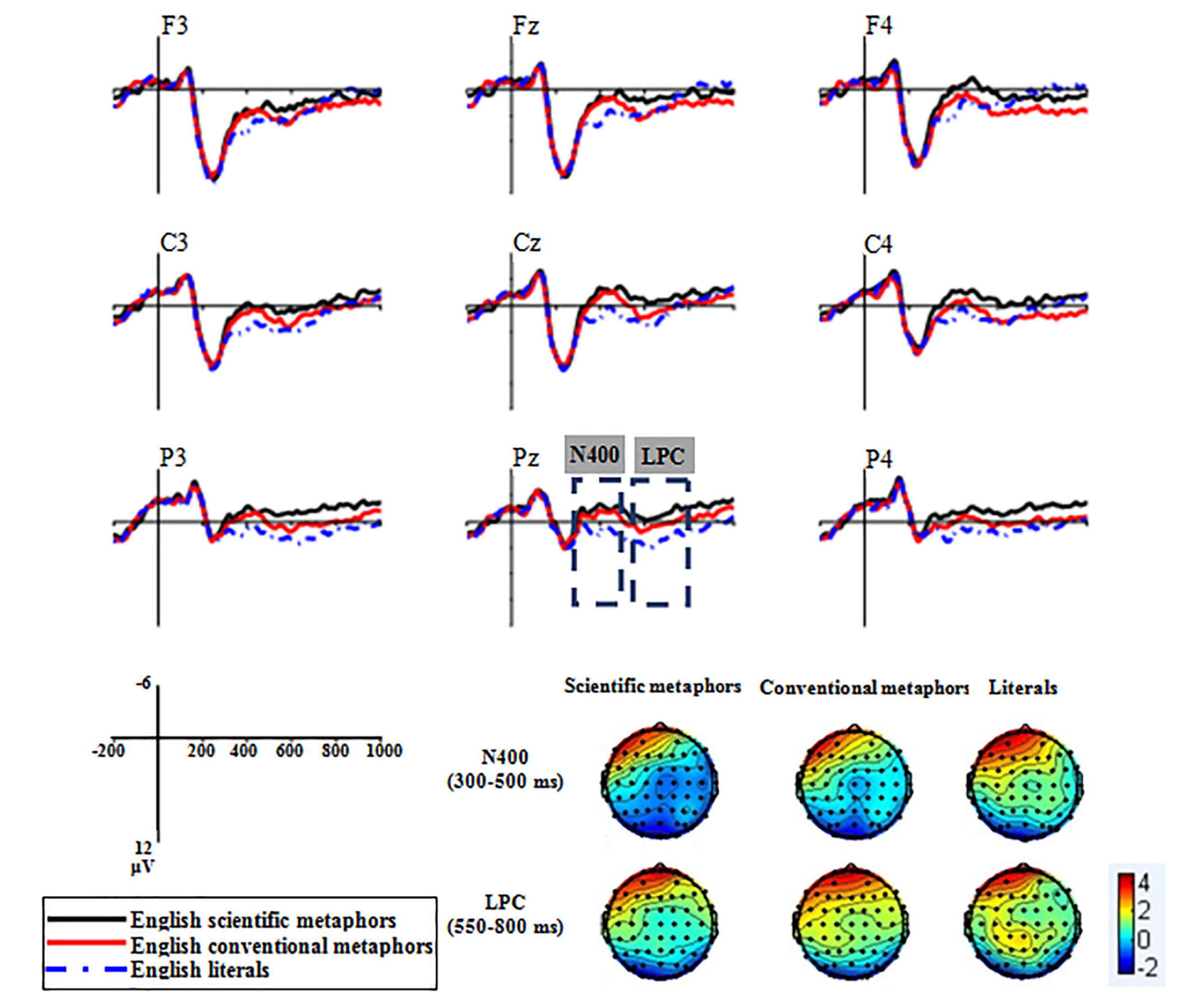

FIGURE 5 | Grand average ERP waveforms of Experiment 2 recorded at the nine chosen electrodes.

[$F(2,32) = 9.78$, $p = 0.001$, $\eta_p^2 = 0.38$]. *Post-hoc* analysis showed that conventional metaphors elicited significantly lower LPCs at the parietal region [$t(16) = -2.92$, $p = 0.01$].

Since between the two metaphorical conditions, the main effect of condition was significant at the parietal region [$F(1,16) = 4.59$, $p = 0.048$, $\eta_p^2 = 0.22$], separate pairwise ANOVAs for the three electrodes at the parietal region (P3, Pz, P4) were performed. It was evident that the LPCs elicited by scientific metaphors were significantly less positive than that of conventional metaphors at P4 [$t(16) = -2.46$, $p = 0.026$] and marginally less significantly positive than that of conventional metaphors at P3 [$t(16) = -1.83$, $p = 0.085$], whereas no such difference was observed at Pz [$t(16) = -1.62$, $p = 0.125$].

Late Negativity. Analogous to Chinese metaphor processing, English (L2) scientific metaphors and conventional metaphors also elicited a late negativity in the LPC time window (see **Figure 6**). The 2 × 3 × 3 ANOVA revealed a marginally significant main effect of condition [$F(1,16) = 3.83$, $p = 0.068$, $\eta_p^2 = 0.19$]. Scientific metaphors elicited a larger late negativity ($M = -1.62$, $SD = 2.46$) than conventional metaphors ($M = -0.46$, $SD = 2.68$). Follow-up analysis showed that the significant effect was only found at the parietal region [$F(1,16) = 4.59$, $p = 0.048$, $\eta_p^2 = 0.22$].

Comparative Results of Experiments 1 and 2

300–500 ms (N400)

In comparing the results manifested by Chinese and English scientific metaphors (see **Figure 7**), a condition × region × hemisphere ANOVA showed that the main effect of condition was not significant [$F(1,16) = 1.79$, $p = 0.46$, $\eta_p^2 = 0.04$], but there were marginally significant

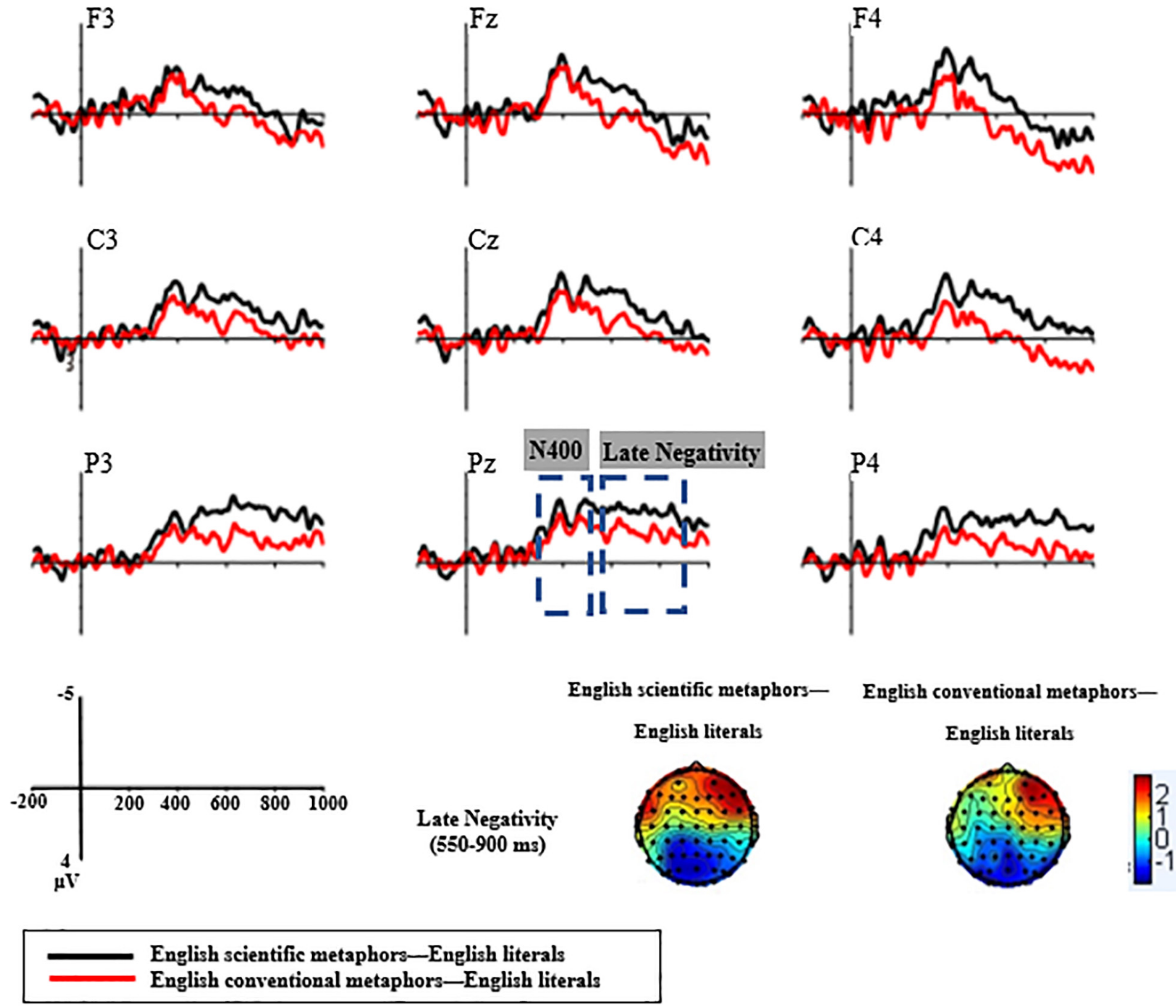

FIGURE 6 | Topographic plots of the two kinds of English metaphors subtracting the literal from the metaphoric in the 300–500 and 550–800 ms time windows.

condition × region interactions [$F(2,32) = 3.37$, $p = 0.073$, $\eta_p^2 = 0.17$]. Pairwise comparisons showed that English scientific metaphors elicited a significantly larger N400 than Chinese scientific metaphors at the parietal region [$t(16) = -2.25$, $p = 0.039$]. There were no significant effects either between Chinese and English conventional metaphors [$t(16) = -0.96$, $p = 0.351$] nor between Chinese and English literal sentences [$t(16) = -0.66$, $p = 0.516$].

550–800 ms (Late Positive Component)

For the difference between Chinese and English scientific metaphors (see **Figure 7**), a condition × region × hemisphere ANOVA showed that, despite the insignificant main effect of condition [$F(1,16) = 1.79$, $p = 0.2$, $\eta_p^2 = 0.10$], significant condition × region interactions were found [$F(2,32) = 9.32$, $p = 0.004$, $\eta_p^2 = 0.37$]. Pairwise comparisons showed that English scientific metaphors elicited a significantly lower LPC than Chinese scientific metaphors at the parietal region [$t(16) = -3.72$, $p = 0.002$], and the difference between the two was highly significant at the right parietal region [P4: $t(16) = -3.99$, $p = 0.001$].

For the difference between Chinese and English conventional metaphors, the condition × region × hemisphere ANOVA revealed a marginally significant main effect of condition [$F(1,16) = 3.08$, $p = 0.098$, $\eta_p^2 = 0.1$] and significant condition × region interactions [$F(2,32) = 6.4$, $p = 0.016$, $\eta_p^2 = 0.29$]. Pairwise comparisons showed that English conventional metaphors elicited a significantly lower LPC than Chinese conventional metaphors at the parietal region [$t(16) = -4.97$, $p < 0.001$], and the difference between two conditions was highly significant at the right parietal region [P4: $t(16) = -5.23$, $p < 0.001$].

In terms of the differences between Chinese and English literal sentences (see **Figure 7**), the condition × region × hemisphere ANOVA revealed a marginally significant main effect of condition [$F(1,16) = 3.54$, $p = 0.078$, $\eta_p^2 = 0.18$] and significant condition × region interactions [$F(2,32) = 3.96$, $p = 0.042$, $\eta_p^2 = 0.2$]. Pairwise comparisons showed that English literal

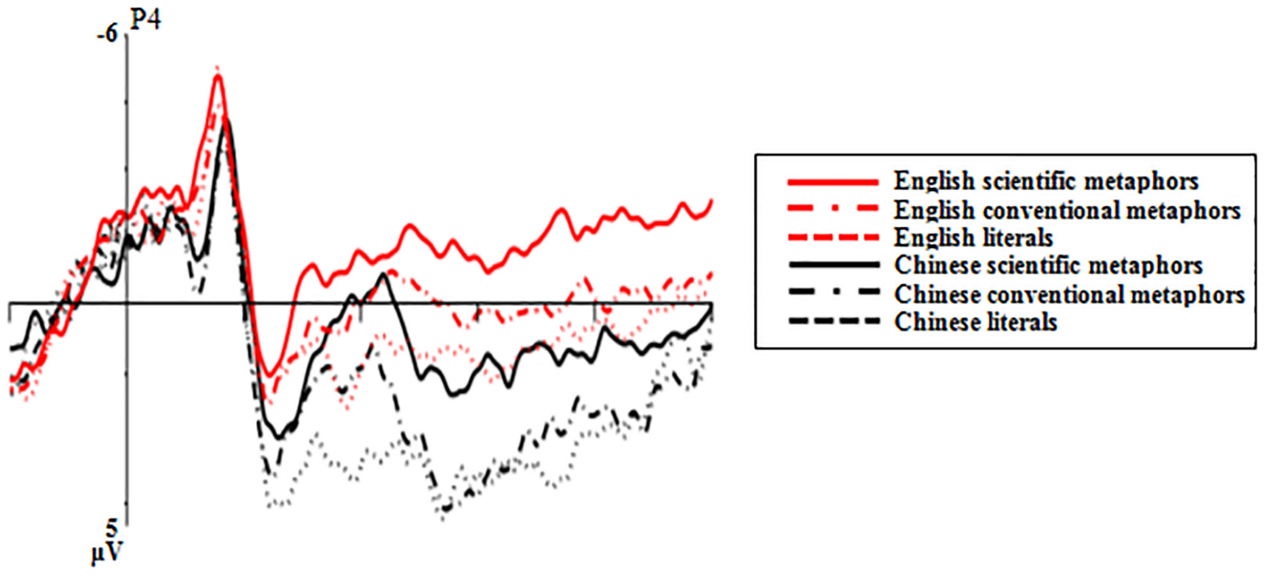

FIGURE 7 | Grand average ERP waveforms of Chinese and English stimuli recorded at the P4 electrode position.

sentences elicited a lower LPC than Chinese literal sentences at the parietal region [$t(16) = -3.59$, $p = 0.002$], and the effect was highly significant at the right parietal region [P4: $t(16) = -5.36$, $p < 0.001$].

Discussion

Taken together, reduced LPC amplitudes for all three English conditions were maximal over the parietal region, especially over the right parietal region, suggesting that the meanings of both L2 metaphoric and literal sentences are integrated with increased cognitive effort. This result might indicate that late speakers who master two languages asymmetrically are less sensitive to levels of conventionality of metaphoric meanings at the later stage of metaphoric language processing (Jankowiak et al., 2017). Moreover, LPCs are typically considered to reflect the depth of syntactic processing. Compared with L2 processing, the understanding of L1 might involve more syntactic analysis (Van Der Meij et al., 2011).

GENERAL DISCUSSION

The first aim of the current study was to observe brain responses to scientific metaphoric, conventional metaphoric, and literal sentences in Chinese (L1) and English (L2). In line with what was hypothesized, in both the Chinese and English experiments, we observed general between-condition differences, with scientific metaphors eliciting higher N400s than conventional metaphors. However, in the English experiment, more significant effects were found in larger regions over frontal, central, and parietal regions, with a slight right hemisphere bias. Importantly, this finding accords with the behavioral results drawn from the Chinese experiment, which showed lower accuracy rates and longer reaction time for scientific metaphors than with conventional metaphors. This supports the graded salience hypothesis for metaphorical expression. The unique complexity and abstraction of scientific language reduce the explicitness of language expression and have an impact on semantic integration in the later stage of processing. The longer distance between the scientific target domain and the daily source domain for scientific metaphor comprehension might increase cognitive effort in conceptual integration (Tang et al., 2017a,b). According to the Career of Metaphor Model (Bowdle and Gentner, 2005), the processing mechanism of metaphors is regulated by their familiarity. The understanding of metaphors with low familiarity requires the construction of metaphorical meaning through lexical semantic processes, while the understanding of metaphors with high familiarity is mainly carried out through semantic retrieval (Coulson and Van Petten, 2002; Lai and Curran, 2013). In addition, compared with literal expressions, conventional metaphors elicited a larger N400 response, which might indicate that despite being frequently used with high familiarity, conventional metaphors require more resource-intensive mappings between concepts than literal expressions.

In both Chinese and English experiments, LPC amplitudes evoked by scientific metaphors were smaller than those elicited by conventional metaphors and literal sentences. Lower LPC amplitudes for scientific metaphors might be caused by a sustained negativity, indexing the integration process of conceptually taxing meanings. Another explanation could be due to reprocessing operations after an initial failure in meaning interpretation (Ruchkin et al., 1988; Rataj, 2014) or alternatively extra working memory load for complex semantic processing (Ruchkin et al., 1988; Anderson et al., 1996; Oberauer et al., 2001; Jiang et al., 2009). According to some monolingual research on metaphors, novel metaphors elicited larger late negativity amplitudes than conventional metaphors, which was interpreted as the continuation of information retrieval or access to the non-literal route when understanding novel metaphors (Arzouan et al., 2007a), and as the ongoing difficulty of meaning integration indexed as the continuation of the N400 effect (Rutter et al., 2012). The larger late negativity of scientific metaphors might further indicate more difficult reintegration of the two domains, especially when the knowledge inference is involved at the later stage of scientific metaphor processing.

Second, the current study aimed to observe whether language nativeness modulates meaning integration mechanisms of scientific metaphors, to which the N400 response is sensitive. More specifically speaking, we were interested in whether L2 scientific metaphors would elicit similar N400s as L1 equivalents for late unbalanced bilingual speakers.

Between-language effects were found with English scientific metaphors eliciting more negative N400s than Chinese ones at the parietal region. Such enhanced N400 response to L2 relative to L1 stimuli is contrary to the results of some bilingual research (Proverbio et al., 2002; Moreno et al., 2008; Midgley et al., 2009; Newman et al., 2012; Heidlmayr et al., 2015). Within the memory system, weaker semantic interconnectivity for L2 compared with L1 words might explain the increased N400 amplitude (Midgley et al., 2009). From the opposite perspective, within the semantic network, larger interconnectivity for L1 words might be linked to the N400 L1/L2 effect (increased N400 amplitudes for L1 relative to L2 words). However, for scientific words, the picture might be different. Compared with conventional words, the semantic connectivity of scientific words might be quite weak in both native and non-native languages. This conjecture would seem to be supported by the behavioral results of the present English experiment. Such weak interconnectivity for scientific words might evoke similar activity in long-term memory in both languages for information retrieval. Unlike the daily words used in conventional metaphors, the scientific terms used in scientific metaphors are even less frequently used, which might result in the much lower familiarity for L2 relative to L1 scientific words. Therefore, the semantic processing of second language scientific terms is more challenging than that of Chinese ones. Moreover, for ordinary L2 speakers (Chang and Wang, 2016), the processing of L2 vocabulary often requires a greater degree of suppression of native vocabulary (Heidlmayr et al., 2015; Wu and Thierry, 2017), which further leads to difficulty in processing L2 words, especially scientific terms.

In addition, at the frontal, central, and parietal regions of the right hemisphere, English scientific metaphors elicited higher N400s than English conventional metaphors, indicating the special role of the right hemisphere in second language processing (Van Der Meij et al., 2011), which supports the Fine-Coarse Semantic Coding Theory (Mashal et al., 2015). Meanwhile, at the left parietal region, there was a marginally significant difference between the two metaphorical sentences, probably showing that the left hemisphere is also involved in understanding L2 scientific metaphors (Kim et al., 2017; Segal and Gollan, 2018).

Third, the current study aimed to observe the modulation of language nativeness for cognitive mechanisms involved at the later stage of meaning reintegration, which the LPC is sensitive to. In line with our predictions, within the LPC time frame, at the parietal region, we observed between-language differences with a slight right hemisphere bias, with smaller LPC responses evoked by the three English conditions than their Chinese counterparts. Reduced LPC amplitudes for scientific and conventional metaphors in the non-native language suggest more demanding cognitive effort to integrate the meaning of both novel and familiar metaphors in an L2 context. This result might indicate that late speakers who master two languages asymmetrically are less sensitive to the levels of conventionality of metaphoric meanings at the later stage of metaphoric language processing.

Inconsistent with the processing of Chinese scientific metaphors, English scientific metaphors only elicited lower LPCs than English conventional metaphors at the parietal region. That is to say, the LPC distribution of L1 scientific metaphors covered a larger area than that of L2 scientific metaphors (Jankowiak et al., 2017; Segal and Gollan, 2018), probably because it might be quite difficult for late L2 learners to reach a similar processing depth as displayed by the native speakers (Newman et al., 2012). In addition, both the left and right hemispheres are involved in the processing of L2 scientific metaphors (Chen et al., 2013), and the right parietal region might play a particularly important role.

CONCLUSION

Through a comparative analysis of the ERP components elicited by scientific metaphors in English and Chinese, this study examined brain responses to scientific metaphors in L2. It was found that nativeness modulates the cognitive cost for semantic integration at an early stage and for semantic reintegration and knowledge inference at a later period, supporting the Career of Metaphor Model and the Graded Salience Hypothesis.

In addition, the scalp distributions of the N400s and the late component elicited by scientific metaphors in L2 reinforce the essential role of the parietal region (especially the right parietal region) in processing L2, supporting the Fine-Coarse Semantic Coding Theory. More abundant types of stimuli could be added to subsequent follow-up experiments for more comparative analysis so as to continuously verify and improve existing studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee in Shaanxi Normal University.

The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XT and WR contributed to conception and design of the study. XT, MH, YH, and SH performed the data collection. LS, MH, and PY performed the analysis. XT wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was funded by National Social Science Foundation of China (Grant No. 17BYY092).

ACKNOWLEDGMENTS

Special thanks to Michael, a British colleague at AHPU, for help to polish some of the language in our final draft.

REFERENCES

Anderson, J. R., Reder, L. M., and Lebiere, C. (1996). Working memory: activation limitations on retrieval. *Cogn. Psychol.* 30, 221–256. doi: 10.1006/cogp.1996.0007

Arzouan, Y., Goldstein, A., and Faust, M. (2007a). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Res.* 1160, 69–81. doi: 10.1016/j.brainres.2007.05.034

Arzouan, Y., Goldstein, A., and Faust, M. (2007b). Dynamics of hemispheric activity during metaphor comprehension: electrophysiological measures. *Neuroimage* 36, 222–231. doi: 10.1016/j.neuroimage.2007.02.015

Bowdle, B. F., and Gentner, D. (2005). The career of metaphor. *Psychol. Rev.* 112, 193–216. doi: 10.1037/0033-295X.112.1.193

Braunstein, V., Ischebeck, A., Brunner, C., Grabner, R. H., Stamenov, M., and Neuper, C. (2012). Investigating the influence of proficiency on semantic processing in bilinguals: an ERP and ERD/S analysis. *Acta Neurobiol. Exp.* 72, 421–438.

Brouwer, H., Fitz, H., and Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Res.* 1446, 127–143. doi: 10.1016/j.brainres.2012.01.055

Chang, X., and Wang, P. (2016). Influence of second language proficiency and syntactic structure similarities on the sensitivity and processing of english passive sentence in late chinese-english bilinguists: an ERP study. *J. Psychol. Res.* 45, 85–101. doi: 10.1007/s10936-014-9319-1

Chen, H., Peng, X., and Zhao, Y. (2013). An ERP study on metaphor comprehension in the bilingual brain. *Chin. J. Appl. Ling.* 36, 505–519. doi: 10.1515/cjal-2013-0034

Coulson, S., and Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Memory Cogn.* 30, 958–968. doi: 10.3758/BF03195780

Coulson, S., and Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Res.* 1146, 128–145. doi: 10.1016/j.brainres.2007.03.008

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., and Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia* 48, 1965–1984. doi: 10.1016/j.neuropsychologia.2010.03.017

Forgács, B., Bardolph, M., Amsel, B. D., DeLong, K. A., and Kutas, M. (2015). Metaphors are physical and abstract: ERPs to metaphorically modified nouns resemble ERPs to abstract language. *Front. Hum. Neurosci.* 9:28. doi: 10.3389/fnhum.2015.00028

Forgács, B., Lukács, Á, and Pléh, C. (2014). Lateralized processing of novel metaphors: disentangling figurativeness and novelty. *Neuropsychologia* 56, 101–109. doi: 10.1016/j.neuropsychologia.2014.01.003

Goldstein, A., Arzouan, Y., and Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain Lang.* 123, 137–142. doi: 10.1016/j.bandl.2012.09.008

Gratton, G., Coles, M. G., and Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr. clin. neurophysiol.* 55, 468–484. doi: 10.1016/0013-4694(83)90135-9

Heidlmayr, K., Hemforth, B., Moutier, S., and Isel, F. (2015). Neurodynamics of executive control processes in bilinguals: evidence from ERP and source reconstruction analyses. *Front. Psychol.* 6:1–17. doi: 10.3389/fpsyg.2015.00821

Jankowiak, K., Rataj, K., and Naskrecki, R. (2017). To electrify bilingualism: electrophysiological insights into bilingual metaphor comprehension. *PLoS One* 12:1–30. doi: 10.1371/journal.pone.0175578

Jiang, X., Tan, Y., and Zhou, X. (2009). Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia* 47, 1799–1815. doi: 10.1016/j.neuropsychologia.2009.02.020

Kaan, E., Harris, A., Gibson, E., and Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Lang. Cogn. Proc.* 15, 159–201. doi: 10.1080/016909600386084

Kim, S. Y., Liu, L., and Cao, F. (2017). How does first language (L1) influence second language (L2) reading in the brain? Evidence from korean-english and chinese-english bilinguals. *Brain Lang.* 171, 1–13. doi: 10.1016/j.bandl.2017.04.003

Kotz, S. A., Rothermich, K., and Schmidt-Kassow, M. (2012). "Sentence comprehension in healthy and braindamaged populations," in *The Handbook of the Neuropsychology of Language* ed. M. Faust (Malden: Blackwell Publishing), 760–777.

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657

Lai, V. T., and Curran, T. (2013). ERP evidence for conceptual mappings and comparison processes during the comprehension of conventional and novel metaphors. *Brain Lang.* 127, 484–496. doi: 10.1016/j.bandl.2013.09.010

Lai, V. T., Curran, T., and Menn, L. (2009). Comprehending conventional and novel metaphors: AN ERP study. *Brain Res.* 1284, 145–155. doi: 10.1016/j.brainres.2009.05.088

Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8, 213. doi: 10.3389/fnhum.2014.00213

Ma, Q., Hu, L., Xiao, C., Bian, J., Jin, J., and Wang, Q. (2016). Neural correlates of multimodal metaphor comprehension: evidence from event-related potentials and time-frequency decompositions. *Int. J. Psychophysiol.* 109, 81–91. doi: 10.1016/j.ijpsycho.2016.09.007

Mashal, N., Borodkin, K., Maliniak, O., and Faust, M. (2015). Hemispheric involvement in native and non-native comprehension of conventional metaphors. *J. Neurolinguist.* 35, 96–108. doi: 10.1016/j.jneuroling.2015.04.001

Midgley, K. J., Holcomb, P. J., and Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *J. Neurolinguist.* 22, 281–300. doi: 10.1016/j.jneuroling.2008.08.001

Moreno, E. M., Fornells, A. R., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguist.* 21, 477–508. doi: 10.1016/j.jneuroling.2008.01.003

Moreno, E. M., and Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish–english bilinguals. *Cogn. Brain Res.* 22, 205–220. doi: 10.1016/j.cogbrainres.2004.08.010

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., and Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of english. *J. Cogn. Neurosci.* 24, 1205–1223. doi: 10.1162/jocn_a_00143

Oberauer, K., Demmrich, A., Mayr, U., and Kliegl, R. (2001). Dissociating retention and access in working memory: an age-comparative study of mental arithmetic. *Memory Cogn.* 29, 18–33. doi: 10.3758/BF03195737

Phillips, N. A., Segalowitz, N., O'Brien, I., and Yamasaki, N. (2004). Semantic priming in a first and second language: evidence from reaction time variability and event-related brain potentials. *J. Neurolinguist.* 17, 237–262. doi: 10.1016/s0911-6044(03)00055-1

Proverbio, A. M., Cok, B., and Zani, A. (2002). Electrophysiological measures of language processing in bilinguals. *J. Cogn. Neurosci.* 14, 994–1017. doi: 10.1162/089892902320474463

Pynte, J., Besson, M., Robichon, F. H., and Poli, J. (1996). The time-course of metaphor comprehension: an event-related potential study. *Brain Lang.* 55, 293–316. doi: 10.1006/brln.1996.0107

Rataj, K. (2014). Surfing the brainwaves of metaphor comprehension. *Poznan Stud. Contemp. Ling.* 50, 55–73. doi: 10.1515/psicl-2014-0004

Ruchkin, D. S. Mahaffey, D., and Sutton, S. (1988). Toward a functional categorization of slow waves. *Psychophysiology* 25, 339–353. doi: 10.1111/j.1469-8986.1988.tb01253.x

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., and Abraham, A. (2012). Can clouds dance? Part 2: an ERP investigation of passive conceptual expansion. *Brain Cogn.* 80, 301–310. doi: 10.1016/j.bandc.2012.08.003

Segal, D., and Gollan, T. H. (2018). What's left for balanced bilinguals? Language proficiency and item familiarity affect left-hemisphere specialization in metaphor processing. *Neuropsychology* 32, 866–879. doi: 10.1037/neu0000467

Steinhauer, K., Drury, J. E., Portner, P., Walenski, M., and Ullman, M. T. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia* 48, 1525–1542. doi: 10.1016/j.neuropsychologia.2010.01.013

Tang, X., Qi, S., Jia, X., Wang, B., and Ren, W. (2017a). Comprehension of scientific metaphors: complementary processes revealed by ERP. *J. Neurolinguist.* 42, 12–22. doi: 10.1016/j.jneuroling.2016.11.003

Tang, X., Qi, S., Wang, B., Jia, X., and Ren, W. (2017b). The temporal dynamics underlying the comprehension of scientific metaphors and poetic metaphors. *Brain Res.* 1655, 33–40. doi: 10.1016/j.brainres.2016.11.005

Tartter, V. C., Gomes, H., Dubrovsky, B., Molholm, S., and Stewart, R. V. (2002). Novel metaphors appear anomalous at least momentarily: evidence from N400. *Brain Lang.* 80, 488–509. doi: 10.1006/brln.2001.2610

Van Der Meij, M., Cuetos, F., Carreiras, M., and Barber, H. A. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology* 48, 44–54. doi: 10.1111/j.1469-8986.2010.01039.x

van Herten, M., Kolk, H. H. J., and Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cogn. Brain Res.* 22, 241–255. doi: 10.1016/j.cogbrainres.2004.09.002

Weber-Fox, C. M., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256. doi: 10.1162/jocn.1996.8.3.231

Weiland, H., Bambini, V., and Schumacher, P. B. (2014). The role of literal meaning in figurative language comprehension: evidence from masked priming ERP. *Front. Hum. Neurosci.* 8:1–17. doi: 10.3389/fnhum.2014.00583

Wu, Y. J., and Thierry, G. (2017). Brain potentials predict language selection before speech onset in bilinguals. *Brain Lang.* 171, 23–30. doi: 10.1016/j.bandl.2017.04.002

Zhao, M., Meng, H., Xu, Z., Du, F., Liu, T., Li, Y., et al. (2011). The neuromechanism underlying verbal analogical reasoning of metaphorical relations: an event-related potentials study. *Brain Res.* 1425, 62–74. doi: 10.1016/j.brainres.2011.09.041

Zucker, L., and Mudrik, L. (2019). Understanding associative vs. abstract pictorial relations: an ERP study. *Neuropsychologia* 133, 1–16. doi: 10.1016/j.neuropsychologia.2019.107127

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 13 July 2022
doi: 10.3389/fpsyg.2022.877997

Check for updates

Is the Processing of Chinese Verbal Metaphors Simulated or Abstracted? Evidence From an ERP Study

*Ying Li[1], Xiaoxiao Lu[1,2], Yizhen Wang[3], Hanlin Wang[4] and Yue Wang[1]**

[1]School of Education, Zhengzhou University, Zhengzhou, China, [2]School of Psychology, Central China Normal University, Wuhan, China, [3]School of International Studies, Zhengzhou University, Zhengzhou, China, [4]Department of Psychology, Hebei Normal University, Shijiazhuang, China

The theory of embodied semantics holds that verbal metaphors are strongly grounded in sensorimotor experience. Many studies have proven that besides sensorimotor simulation, the comprehension of verbal metaphors also requires semantic abstraction. But the interaction between simulation and abstraction, as well as the time course of metaphorical meaning integration, is not well understood. In the present study, we aimed to investigate whether embodiment or abstraction, or both, is employed in the processing of Chinese verbal metaphor. Participants were asked to read subject-verb metaphorical, verb-object metaphorical, literal-concrete and literal-abstract sentences, and the target words were measured at the verb and the object of each sentence. The results revealed that a similar N400 effect was elicited by the target verbs in the verb-object metaphorical and the literal-concrete sentences, and a similar P600/LPC effect was induced by the target verbs in the subject-verb metaphorical and the literal-abstract sentences, reflecting that the verb-object metaphors trigger a simulation process, while the subject-verb metaphors trigger an abstraction process in the verb processing stage. Moreover, the subject-verb metaphors elicited a stronger P600/LPC effect by the target verbs than the verb-object metaphors, but there was no difference of the P600/LPC caused by the target objects between the two kinds of metaphors, revealing that the metaphorical meaning of a subject-verb metaphor is integrated in the verb processing stage, while that of a verb-object metaphor is reanalyzed in the object processing stage. These results suggest that a verbal metaphor is processed both by simulation and abstraction, and the metaphorical meaning is integrated immediately with the unfolding of the sentence meaning. The position where the semantic conflict lies in a sentence (verb vs. object) modulates the time course of metaphor sentence comprehension.

Keywords: verbal metaphor, simulation, abstraction, N400, P600/LPC

OPEN ACCESS

Edited by:
Huili Wang,
Dalian University of Technology, China

Reviewed by:
Shuo Cao,
Dalian University of Technology, China
Lin Chen,
University of Pittsburgh, United States

****Correspondence:***
Yue Wang
yuezi68@126.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *17 February 2022*
Accepted: *13 June 2022*
Published: *13 July 2022*

Citation:
Li Y, Lu X, Wang Y, Wang H and Wang Y (2022) Is the Processing of Chinese Verbal Metaphors Simulated or Abstracted? Evidence From an ERP Study. Front. Psychol. 13:877997. doi: 10.3389/fpsyg.2022.877997

INTRODUCTION

A metaphor refers to using the flexibility of semantic features to express some new meanings by forming semantic conflict with its literal meaning (Semino et al., 2008; Rutter et al., 2012; Benedek et al., 2014). Unlike literal languages, in metaphorical expressions, when a sentence produces such a semantic conflict, the novel meaning created is not inappropriate but conveys a figurative, abstract sense. Dedre Gentner and France (1988) suggested that verbal metaphors would be generated when the collocation of a verb and a noun is unconventional. For example, in the expression "The rumor flew through the office," the subject of the action verb "fly" is an abstract and inanimate agent "rumor" which cannot physically perform the action of "fly." Therefore, the agent (e.g., rumor) and the verb (e.g., fly) constitute a strong conflict, producing a metaphorical expression that means "rumor spreads very fast." Similarly, it is clear that when someone says "he catapulted his words from the dais," the speaker does not literally mean that the orator uses a catapult to bombard the audience, but rather that s/he speaks with vehemence (Obert et al., 2018). Verbal metaphor is not only frequently used in daily life and literary works (Cameron and Gibbs, 2008), but also leads to a wide range of scientific research on how action-based metaphors are comprehended.

A verbal metaphor is a linguistic construction that exemplifies the embodied nature of cognition as verbs generally entail more action content (Feng and Zhou, 2021). The embodied semantic view of metaphors holds that the processing of verbal metaphors relies on sensorimotor simulation, as the verbal metaphors are grounded in the physical body and sensorimotor system (Barsalou et al., 2003; Gibbs, 2006; Qu et al., 2013; Yin et al., 2013). From this perspective, to understand a verbal metaphor such as "The media bent the truth," we need to simulate the concrete act of "causing to curve." Wilson and Gibbs (2007) found that people are faster to assess whether or not a sentence is meaningful if they perform or imagine performing a congruent motion before reading a metaphor (e.g., a grasp motion before reading "grasp a concept"), suggesting that comprehension of the metaphorical term "grasp" draws from simulation of its literal meaning. The embodied simulation view has been supported recently by neural studies. It has been demonstrated that there are sensorimotor activations across brain regions when participants read verbal metaphors (Desai et al., 2011; Boulenger et al., 2012; Lauro et al., 2013; Lai et al., 2019). For example, reading verbal metaphors related to motion content (e.g., grasping a concept) can activate brain regions involved in motor perception and planning associated with hands (Desai et al., 2013; Lauro et al., 2013). A few secondary motor regions are found to be involved when participants read familiar verbal metaphors, while the primary sensory and motor regions are more active when they read unfamiliar novel metaphors (Desai et al., 2011, 2013; Cardillo et al., 2012), suggesting that novel predicate metaphors rely more on sensorimotor information corresponding to the verbs. To sum up, these studies support the view of embodied simulation that verbal metaphor comprehension counts on sensorimotor simulation, and reading a verbal metaphor activates the sensorimotor system associated with the embodied experiencing of concrete conceptual domains of the metaphor.

However, in the psycholinguistic literature on this topic, the dominant assumption is that metaphorical representation is processed as abstractions rather than concrete representations. Abstractions are conceptual representations that are less specified (Gentner and Asmuth, 2019) than their literal-concrete counterparts. According to this view, verbal metaphor comprehension is a result of abstraction processing of the semantic system, where literal-level concrete features of a concept play a little-to-no role in metaphor comprehension. Such as, when reading the familiar expression "The media bent the truth," one would directly retrieve the abstract meaning "distort" from long-term memory. Compared with literal sentences that convey physical senses (e.g., The repairman bent the pipe), verbal metaphors use verbs figuratively *via* abstraction from concrete action terms. Desai et al. (2011) found that the activation of sensorimotor regions decreases with the increase of sentence abstraction. Raposo et al. (2009) reported more specifically that when verbs are presented in literal-concrete sentences, there is greater activation in the sensorimotor areas and frontotemporal lobes associated with language processing. However, when verbs are presented in metaphorical sentences, there is no activation in the motor and premotor areas. Meanwhile, Chatterjee (2010) also proved that understanding a literal sentence with an action verb activates the left occipital and temporal motor area. Instead, when reading a metaphorical sentence with the same verb, the inferior frontal gyrus and left temporal gyrus relating to language are more active, indicating that the semantic abstraction system is more involved when dealing with verbs in metaphors.

Although the activation of abstract meaning in metaphorical processing has been reported in a few studies, it is unclear how the abstract sense and the concrete one interact and work together to access the metaphorical meaning. If the comprehension of verbal metaphors is to combine the neural patterns of literal-concrete meaning processing and literal-abstract one, it is pressing to explore at what stage the concrete and abstract meanings in verbal metaphors are activated. In other words, are the verbal metaphors comprehended in their concrete sense first, or can metaphorical meaning be extracted directly through semantic abstraction? If literal-concrete meaning activation is early, then it would support the critical role of concrete and bodily experiences in comprehending abstract meaning (Gallese and Lakoff, 2005). In contrast, if it is activated late, then such activation can be interpreted as being epiphenomenal, suggesting that metaphorical meaning is accessed directly through semantic abstraction (Mahon and Caramazza, 2008). Therefore, the present study intended to use Event-Related Potentials (ERP) to investigate the activation timings of literal-concrete and literal-abstract meaning in metaphorical sentences, which can reveal whether the processing of verbal metaphors is simulated or abstracted.

Previous research adopted electrophysiological indexes to detect the processing of the verbal metaphors, and highlighted that the N400 may reflect activation of concrete word's multimodal information, which is related to sensorimotor recruitment. For example, Lai et al. (2019) found that when the verb presented, both metaphorical and literal sentences including concrete action verbs elicited larger N400 effect than the sentences containing abstract verbs. Holcomb et al. (1999) suggested that more semantic information is activated in the long-term semantic memory *via* concrete words than abstract words. However, when the semantic association strengths were controlled constant for concrete and abstract words, the concreteness N400 effect was still found in concrete words (Barber et al., 2013). Thus, the concreteness N400 is believed to reflect embodied simulation process with sensory-motor recruitment. In addition to the concreteness N400 effect in metaphor processing, the P600/LPC component is also commonly found to index semantic reanalysis and integration process, with great significance to understand the specific association between P600/LPC and abstract semantics in verb metaphor comprehension. A few studies also found a distinct P600/LPC effect in the metaphor condition relative to the literal condition, suggesting that participants were aware of semantic conflicts 450–750 ms after encountering the unfamiliar metaphorical word and were able to integrate the abstract meaning of the action word into the sentence context *via* semantic reanalysis (Shen et al., 2015; Obert et al., 2018). In addition, these findings also suggest that both the sensorimotor system and the abstract sense processing system are involved in the verbal metaphorical processing, and the integration of metaphorical meaning is based on concrete simulation and abstraction. Based on the previous findings, the current study will further explore the activation time of concrete sense and abstract sense in verbal metaphor processing, which is related to N400 and P600/LPC, respectively.

In addition, previous studies have mainly focused on the timing of sensory-motion system activation in verbal metaphor comprehension (Lai et al., 2019), while the time course of metaphorical meaning integration has been underestimated. As verbal metaphors occur from the semantic conflict with literal meanings between verbs and context, the timing of the conflict is critical to the integration of metaphorical meanings. By manipulating animacy violation of the subject in the verbal metaphorical sentences, the previous study found that the metaphors with inanimate actor elicited an attenuated P600 as compared with the animate counterparts and converged to the same level as literal sentences (Ji et al., 2020). The result indicates that integration of metaphorical meaning is instantaneously accessed with the sentence meaning unfolding, thus the semantic conflict between verbs and different sentence components may lead to different time processes of metaphorical meaning extraction. Specifically, semantic conflict of verbal metaphors may occur between the verb and inanimate actors of the sentence, such as in subject-verb metaphor. It also appears when the object is presented, such as in verb-object metaphors. Therefore, it is to be examined whether the metaphorical meaning of subject-verb metaphors is integrated earlier than the verb-object metaphors. Since previous studies did not make a syntactic distinction between these two metaphors, it remains unclear that whether the semantic reanalysis and integration of abstract meaning are conducted once an action verb appears or it remains unfolded until the end of the sentence.

Taken together, we form the following two questions in the current study. First, in the processing of verbal metaphors, are the literal concrete meanings indicating sensorimotor simulations and/or the abstraction-related meanings activated? Second, at what time are these meanings integrated into the sentence? Answering these questions can further clarify the processing of verbal metaphors. Based on the above considerations, the current study aims to investigate the processing mechanism of verbal metaphors, focusing on the time course of metaphorical meaning integration. The N400 and P600/LPC components are used as indexes for the activation of associated sensorimotor simulation and abstract sense processing system. We assume that besides sensorimotor simulation, the comprehension of verbal metaphors also requires semantic abstraction. According to previous research, the N400 effect induced by verbs in verb-object metaphors will be similar to that in literal-concrete sentences, which recruits more perceptual motion simulation. And the P600/LPC effect elicited by verbs in subject-verb metaphors will be similar to that in literal-abstract sentences, reflecting the semantic aspects of processing, including processing and integration processes of abstract meaning.

Meanwhile, in the experiment, by setting the literal-metaphorical conflict, respectively, at the verb and the object, the timing of generating the metaphorical meaning is going to be distinguished. If the semantic integration of verbal metaphors occurs immediately with the unfolding of sentence meaning, then, participants will reanalyze the subject-verb metaphor sentences by integrating the meaning with the previous subject after the verb is presented, as well as access the appropriate metaphorical meaning of the sentence. In comparison, semantic reanalysis and integration will be triggered only when the object word of the verb-object metaphor is presented at the end of the sentence. Thus, we argue that the P600/LPC effect induced by the action verb tends to be revealed greater in the subject-verb metaphor sentences than that in the verb-object metaphor sentences.

MATERIALS AND METHODS

Participants

A priori power analysis using G*Power 3.1 (Faul et al., 2007) suggested that, for a single-factor within-subject design with a power of 80%, approximately 24 participants for large effect size ($d = 0.8$) would be needed for our experiments. Additionally, we referred to the sample sizes used in previous studies (Lai et al., 2019; Ji et al., 2020), and consequently recruited 40 Chinese students as participants. All participants were completely unaware of the purpose of the present experiment. Written informed consent was obtained before the study. All participants had normal or corrected-to-normal vision, as established through self-report. The data for two participants were discarded from the statistical analysis as a result of excessive electroencephalogram (EEG) artifacts. Data of 38 healthy participants (17 male, 21 female; age 18–26 years; Mage = 20 years) were used for further analysis.

Material

Some of the original materials were collected from the BCC (Beijing Language and Culture University-Corpus Center of China) and CCL (Peking University Modern Chinese Corpus) corpus, and the other part of the original materials are selected from the language of daily life and other genres of literature. The experiment contains four conditions of sentence materials (See **Table 1**): Subject-verb metaphorical sentences (SVM), Verb-object metaphorical sentences (VOM), literal-abstract sentences (LA) and literal-concrete sentences (LC). Subject-verb metaphor (SVM) condition referred to a sentence with an inanimate agent followed by an action verb that the agent could not physically perform. Therefore, the semantic conflict point falls on the verb of the sentence. In the verb-object metaphor (VOM) condition, the subject of the sentence is an animate agent, followed by an action verb and an abstract noun object. Since its object could not be physically manipulated, its conflict point falls on the object of the sentence. In the literal-abstract (LA) condition, the same inanimate agent was used, but the action verb was replaced by an abstract verb with a similar meaning. In the literal-concrete (LC) condition, the subject was an animate agent that could physically manipulate an object while the action verb remained the same as the verb in the metaphorical condition. Some of the sentences from the corpus were adapted for the consistency of SVO (NP1 + V + NP2) syntactic structure across all conditions, controlling the length of the sentence and visual complexity. And there were no statistical difference in words frequency of verb and object between conditions [verbs: (0.003 ± 0.003) vs. (0.004 ± 0.006); object: (0.011 ± 0.015) vs. (0.003 ± 0.006)].

First, a questionnaire was adopted to evaluate and screen materials. The initial stimuli consisted of 80 sets of sentences. Two hundred and forty college students who would not participate in the formal experiment filled in the questionnaire to rate the acceptability, familiarity and comprehensibility of the experimental materials on a 5-point Likert scale (1 = completely unacceptable/highly unfamiliar/highly unintelligible; 5 = completely acceptable/highly familiar/highly intelligible). According to the evaluation results, 25 sets of sentences were selected for the formal experiment. The ratings of the four experimental sentences in acceptability, familiarity and comprehensibility were all above 3. In addition, 25 nonsense sentences were selected as filler materials, and the ratings of filler sentences in all dimensions were below 3.

We recruited 32 participants who neither took part in the evaluation experiment nor in the formal experiment to complete the sentence comprehension task and re-evaluate the experimental materials. The sentences were presented word by word on a computer screen. At the end of each sentence, the participants were asked to judge whether the sentence conveys a reasonable meaning, and press a corresponding key. One-way repeated-measures ANOVA of response time to the sentence comprehension task showed that there was no significant difference on RTs between the four experimental sentences $F(3, 93) = 1.76$, $p = 0.16$. The results suggested that the four experimental sentences were highly acceptable and understandable.

The experimental material consists of 5 lists, with each experimental condition from the same material set assigned to different material lists. Each list contains 25 sentences, of which five sentences per condition. Participants were randomly assigned to one of the lists.

TABLE 1 | Sample stimuli.

Conditions	Examples	Explanation
Subject-verb metaphor (SVM)	公司 / 抓住 / 机会。	The company grasped the opportunity.
Verb-object metaphor (VOM)	老板 / 抓住 / 机会。	The boss grasped the opportunity.
Literal-abstract (LA)	公司 / 获得 / 机会。	The company sought the opportunity.
Literal-concrete (LC)	小哲 / 抓住 / 绳子。	Xiaozhe grasped the rope.

Procedure

Participants were seated in a comfortable chair in a sound-attenuated, electrostatic-shielding room and instructed to read each sentence carefully. In the experiment, each sentence was presented word by word. First, participants performed a practice block of 10 sentences to get familiar with the task and experimental environment. The materials used in the practice stage would not appear in the formal experiment. Each trial began with a fixation point "+" displayed at the center of the screen for 500 ms, and then one sentence was displayed word by word in the center with white characters on a black background. Each word was displayed for 500 ms, with a 500 ms blank screen between words. The last word of the sentence was presented with a period. The participants were instructed to read each sentence silently and attentively. A sentence comprehension question appeared randomly after 40% of the trials during the experiment. The participants were asked to respond "meaningful" or "meaningless" as accurately and quickly as possible by pressing the "F" (meaningful) and "J" (meaningless) keys when "?" was presented, with "F" and "J" being balanced between subjects. If the participant did not press the key for 1,200 ms, the "?" automatically disappeared, and the next stimulus began. The whole experimental process lasted about 1.5 h. The trial structure is illustrated in **Figure 1**.

EEG Recording and Analysis

We recorded EEG data from 64 electrodes (following the international 10–20 system) using a Neuroscan system and referenced to the left and right mastoids, with a ground electrode on the medial frontal aspect. The vertical electrooculogram (EOG) was recorded supra- and infraorbitally at the left eye; the horizontal EOG was recorded from the left versus the right orbital rim. The impedance of all electrodes was maintained at below 5 kΩ. The sampling rate was 1,000 Hz. Data were filtered online using a 0.05–100 Hz band-pass filter.

Offline processing of EEG signal data was performed in MATLAB using EEGLAB and ERPLAB toolbox. EEG data were re-referenced to the average of the left and right mastoids. A digital band-pass filter of 0.01–30 Hz was applied to the EEG recordings. Independent component analysis (ICA) was performed on continuous data for each participant to remove components relevant to eye movements and eye blinks. Epochs ranged from −200 to 800 ms after the onset of the stimulus, with the 200 ms interval preceding the stimulus onset serving

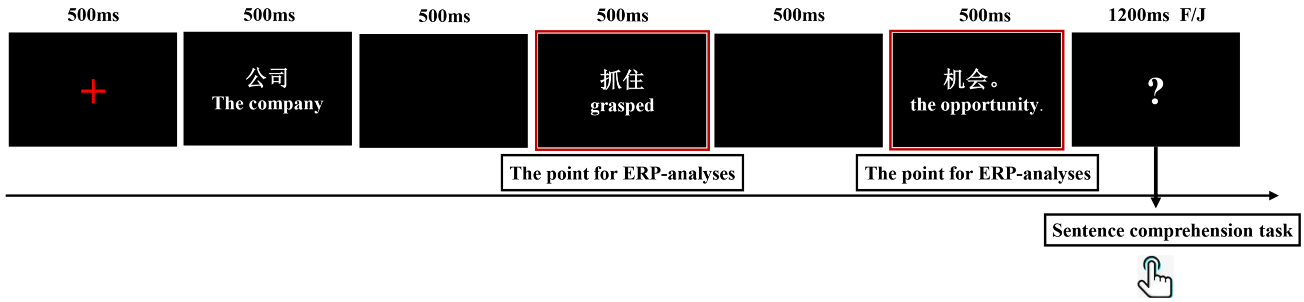

FIGURE 1 | Trial structure of the experiment.

as the baseline. Any epoch with EEG voltages exceeding a threshold of ±100 μV was excluded from the average.

In accordance with previous studies on N400 and P600 for metaphor (Shen et al., 2015; Lai et al., 2019) and the visual inspection based on the grand averaged data, we defined two time windows: 380–500 ms for the N400 and 670–770 ms for the P600 both for verb and object. Values were subjected to a 4 (sentence type: SVM, VOM, LA, LC) × 3 (electrode area: frontal, central, parietal) repeated-measures ANOVA. We classified these electrode sites into three areas: frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal (P3, Pz, P4). All repeated-measures ANOVA results received Greenhouse–Geisser correction if the sphericity assumption was violated. Post-hoc multiple comparisons were carried out using Bonferroni-adjusted corrections. Effect sizes were presented as partial eta-squared (η_p^2) for F tests.

RESULTS

The accuracy of the sentence comprehension questions was 88.5% (SD = 10.5%), indicating that the participants were engaged in reading. The total mean amplitudes for the target verb and object indicated that each condition induced significant N400 and P600/LPC (see **Figures 2, 3**).

Verb Processing Stage

To investigate the processing mechanism and the activation time of concrete or abstract meaning of verbal metaphors, we analyzed the average amplitudes of N400 and P600/LPC detected when the verb appeared. The statistical results are shown in **Table 2**.

N400

A repeated-measures ANOVA revealed significant main effects of both sentence type F (3, 111) = 3.38, $p < 0.05$, $\eta_p^2 = 0.08$, and electrode area F (2, 74) = 63.08, $p < 0.001$, $\eta_p^2 = 0.63$. Further analysis indicated that the waveforms for the verb-object metaphors and literal-concrete sentences were much more negative than those for the subject-verb metaphors and literal-abstract sentences, while there was no significant difference between verb-object metaphors and literal-concrete sentences. The mean amplitude of N400 decreased gradually in the frontal, central and parietal brain regions ($ps < 0.05$). No significant interaction was reported between sentence type and electrode.

P600/LPC

A repeated-measures ANOVA revealed a significant main effect of sentence type F (3, 111) = 3.86, $p < 0.05$, $\eta_p^2 = 0.09$, with the subject-verb metaphors and literal-abstract sentences inducing significantly more positive mean amplitudes than verb-object metaphors, while the mean amplitudes for subject-verb metaphors and literal-abstract conditions showed no significant difference. There was neither a main effect of electrode area nor the interaction between sentence type and electrode area.

Object Processing Stage

To investigate the time course of semantic integration of metaphorical comprehension, we analyzed mean amplitudes of N400 and P600/LPC observed in the object processing stage when the sentence meaning was fully accessible. The statistical results are shown in **Table 3**.

N400

A repeated-measures ANOVA revealed significant main effects of sentences type F (3, 111) = 3.42, $p < 0.05$, $\eta_p^2 = 0.09$, and electrode area F (2, 74) = 27.12, $p < 0.001$, $\eta_p^2 = 0.42$. Pairwise comparisons showed that the literal-concrete condition induced a stronger N400 effect than the other three conditions. The mean amplitude of N400 decreased gradually in the frontal, central and parietal brain regions. There was a significant interaction between sentence type and electrode area F (6, 222) = 3.27, $p < 0.01$, $\eta_p^2 = 0.08$. Simple-effect tests using F-test showed that the literal-concrete condition induced a larger N400 effect than the other three conditions in the frontal and central regions ($ps < 0.05$).

P600/LPC

A statistical analysis revealed a main effect of sentence type F (3, 111) = 3.08, $p < 0.05$, $\eta_p^2 = 0.08$. Pairwise comparisons showed that the waveform for the subject-verb metaphors was much more positive than that for the literal-abstract sentences, and the waveform for the verb-object metaphors

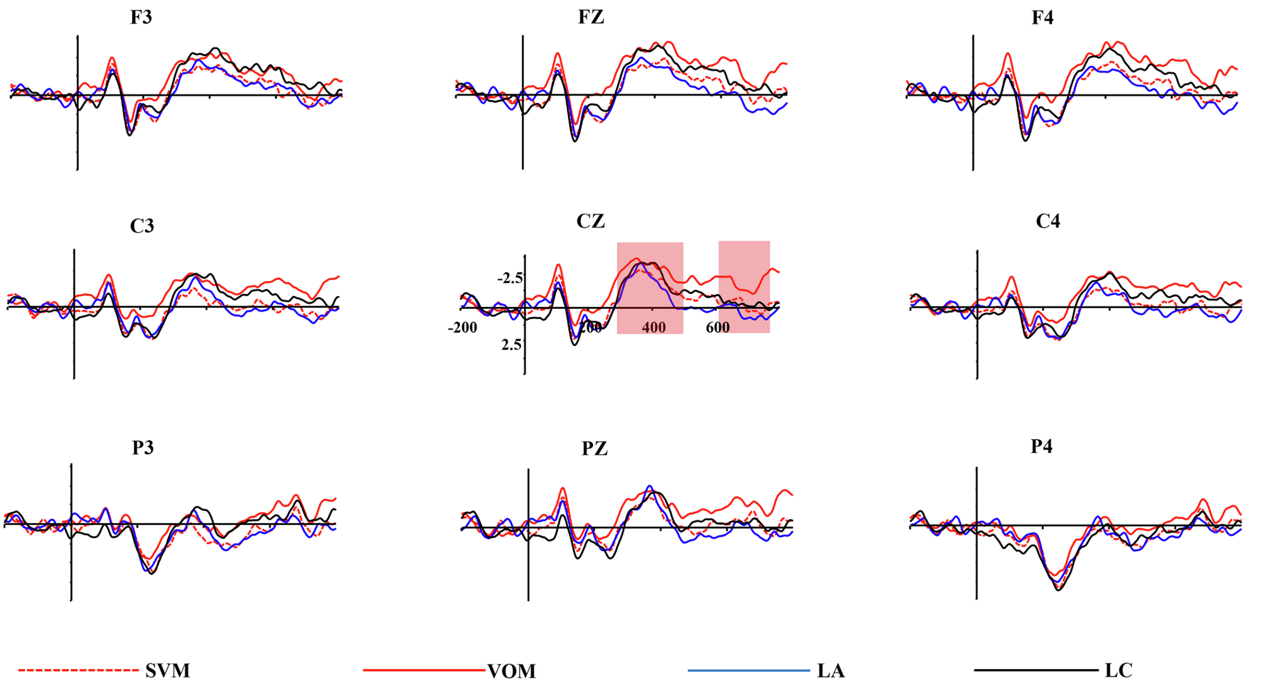

FIGURE 2 | Grand average ERP waveforms recorded at verb for the chosen electrodes.

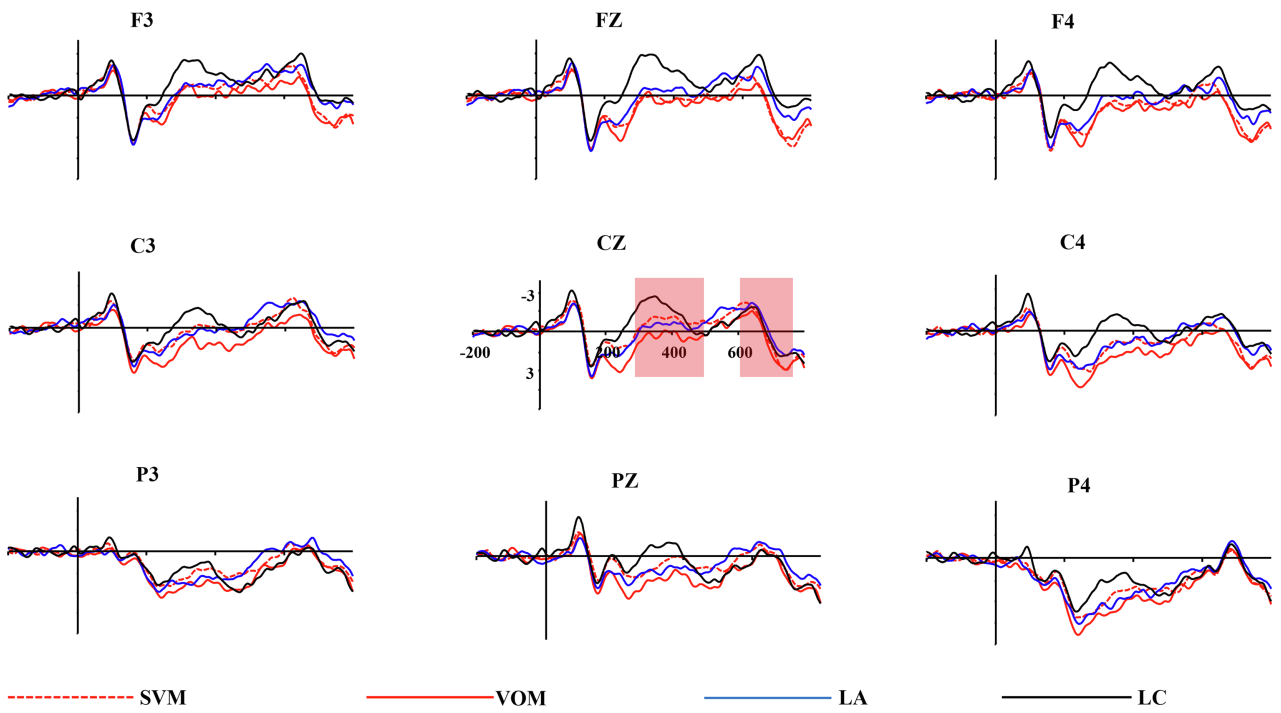

FIGURE 3 | Grand average ERP waveforms recorded at object for the chosen electrodes.

was much more positive than that for the literal-concrete sentences, while there was no significant difference between the subject-verb metaphors and verb-object metaphors. There is an interaction between sentence type and electrode area $F(6, 222) = 5.76$, $p < 0.001$, $\eta_p^2 = 0.14$. Simple-effect tests using F-test showed that the subject-verb metaphors and verb-object metaphors elicited stronger P600/LPC effects than literal-concrete and literal-abstract conditions in the

TABLE 2 | Mean amplitude (μV) and standard deviation (*SD*) of N400 and P600/LPC for target verb.

EEG component	Electrode area	Sentence type			
		SVM	VOM	LA	LC
N400	Frontal	−1.70 ± 0.47	−2.83 ± 0.64	−1.46 ± 0.55	−2.47 ± 0.56
	Central	−0.90 ± 0.51	−2.00 ± 0.58	−0.67 ± 0.51	−1.63 ± 0.53
	Parietal	0.57 ± 0.45	−0.33 ± 0.46	0.55 ± 0.47	−0.09 ± 0.45
P600/LPC	Frontal	0.17 ± 0.34	−0.84 ± 0.60	0.66 ± 0.53	−0.47 ± 0.52
	Central	0.14 ± 0.34	−1.36 ± 0.53	0.57 ± 0.48	−0.39 ± 0.47
	Parietal	−0.15 ± 0.31	−1.22 ± 0.46	0.20 ± 0.44	−0.37 ± 0.45

TABLE 3 | Mean amplitude (μV) and standard deviation (*SD*) of N400 and P600/LPC for target object.

EEG component	Electrode area	Sentence types			
		SVM	VOM	LA	LC
N400	Frontal	0.02 ± 0.62	0.35 ± 0.62	−0.35 ± 0.59	−1.05 ± 0.64
	Central	0.00 ± 0.56	1.03 ± 0.59	0.09 ± 0.56	−0.26 ± 0.58
	Parietal	1.44 ± 0.54	2.46 ± 0.59	1.72 ± 0.54	1.72 ± 0.51
P600/LPC	Frontal	1.75 ± 0.58	1.76 ± 0.53	0.45 ± 0.57	−0.08 ± 0.59
	Central	1.09 ± 0.52	1.31 ± 0.53	0.22 ± 0.48	0.39 ± 0.59
	Parietal	0.76 ± 0.48	0.93 ± 0.49	0.22 ± 0.42	0.80 ± 0.52

frontal region ($ps < 0.05$). There existed no main effect of electrode area.

DISCUSSION

The present study explored the neural mechanism and temporal processing of Chinese verbal metaphors from the sentence level. It was found that a similar N400 effect was elicited by the action verbs in the verb-object metaphorical and the literal-concrete sentences, both inducing more negative N400s than the subject-verb metaphors. Instead, the subject-verb metaphors elicited a more positive P600/LPC than the verb-object metaphors, which was similar to the literal-abstract sentences. When the object word was presented, both types of verbal metaphor induced a stronger P600/LPC effect than the literal sentences. Taken together, these results suggest that the verb-object metaphor activated more concrete meaning, while the subject-verb metaphor activated more literal-abstract meaning when the action word is processed. Moreover, in the final stage of semantic integration, both types of verbal metaphors need to be reanalyzed to ensure that the metaphorical meaning is successfully integrated into the whole sentence.

Activation of Concrete and Abstract Meanings in Verbal Metaphors

The view of embodied semantics holds that verbal metaphors are strongly grounded in sensorimotor experience (Richardson et al., 2003; Gibbs, 2006; Wilson and Gibbs, 2007; Bardolph and Coulson, 2014). Accordingly, a verbal metaphor is processed as a simulation in which the activation of literal concrete meaning plays an important role. While other researchers argue that abstract semantic processing also has a crucial role in verbal metaphor comprehension, in addition to the activation of literal-concrete meaning (Jamrozik et al., 2016; Al-Azary and Katz, 2020). In line with the latter view, our findings found that the verb-object metaphors elicited an N400 effect similar to the literal-concrete condition, while the subject-verb metaphors induced a P600/LPC effect similar to the literal-abstract condition in the verb processing stage, confirming a combination of simulation and abstraction processing in verbal metaphor comprehension.

Recent studies confirm that the involvement of sensorimotor network and abstraction process is dynamic in metaphor comprehension (Al-Azary and Katz, 2020). Based on previous research, the present study further compares the activation of concrete and abstract meanings in the processing of subject-verb metaphors and verb-object metaphors. It shows that the N400 amplitudes induced by verbs in verb-object metaphorical sentences and literal-concrete sentences were greater than that in the subject-verb metaphors and literal-abstract sentences. Although both conditions contain the same action verb, there is no semantic conflict when processing the verb in a verb-object metaphor. Thus, the concrete meaning of the verb is activated in the first place. In contrast, a subject-verb metaphor produces such a conflict when the verb appears, in which the verb needed to be reanalyzed to activate a reasonable abstract meaning, namely, the metaphorical meaning of it. In the verb processing stage, the comprehension of a verb-object metaphorical sentence is more likely to deal with the literal-concrete meaning, relying on sensorimotor simulation, while processing a subject-verb metaphorical sentence activates the abstract meaning more quickly and extracts the metaphorical semantic relation between the subject and the verb.

In this study, by comparing the two types of verbal metaphors and the corresponding processing of literal-concrete and literal-abstract meanings, the results support the simulation-abstraction hybrid view, that is, both simulation and abstraction are viable mechanisms for processing metaphorical meaning (Al-Azary and Katz, 2020). Verbal metaphor comprehension is not entirely based on sensorimotor simulation but also counts on semantic abstraction. More important, the concrete and abstract semantic activation timeframes are regulated by the place where the semantic conflict lies.

The Integration of the Metaphorical Meaning in Verbal Metaphors

The EEG activities in the verb processing phase demonstrated that both concrete and abstract meaning were activated in a verbal metaphor. However, the metaphorical meaning of the whole sentence remains uncertain at that stage, since the syntactic meaning has not been fully integrated yet. Especially for the verb-object metaphorical condition, only when the object appears can it constitute a semantic conflict to create a metaphorical meaning. Therefore, we further analyzed the ERP waveforms elicited by the object word in which the sentence meaning is being completely processed.

The results showed that in the object processing stage, both subject-verb metaphors and verb-object metaphors induced stronger P600 effect than literal sentences. When the object word is presented, the verb and object constitute a semantic conflict in both metaphorical conditions. Therefore, both types of metaphors need to be integrated with the context through the reanalysis of the sentence meaning, costing more cognitive resources than the processing of a literal sentence (Yang et al., 2013). In this phase, a verb-object metaphorical sentence creates a semantic conflict between the verb and the object, and the sentence meaning needs to be re-processed, resulting in a significant P600/LPC effect (Ji et al., 2020). While for a subject-verb metaphor, the abstract meaning of the verb has been activated during this period for the initial integration of metaphorical meaning, and semantic processing will continue with the unfolding of the sentence context until the end of the sentence. The pre-activation of the abstract sense of a subject-verb metaphor weakens the conflict between the verb and the object at the end of the sentence.

In addition, our findings are consistent with recent studies (e.g., Lai et al., 2019). The verb of subject-verb metaphorical sentence induced an N400 effect, and the very same sensory-motor recruitment in concrete literal language takes place for comprehending subject-verb metaphoric expressions. In other words, the specific action meaning of the verb at this stage is also activated, and only when the object finally appears, there shows a semantic conflict to be reintegrated with the previous context, so a P600/LPC effect similar to that induced by the verb-object metaphors is revealed. The P600/LPC effect also reflects that the metaphorical meaning of the sentence can be directly extracted and integrated through abstraction processing. By analyzing the semantic activation and integration at the different time windows of semantic processing, the study shows that the metaphorical meaning in the verbal metaphorical comprehension is instantly integrated and extracted as the meaning of the sentence is unfolded, and the time when semantic conflict appears will affect the activation time of concrete and abstract meanings in the metaphor.

Another interesting finding in this study is that the N400 elicited by literal-concrete condition in the frontal and central regions is larger than the other three conditions at the object processing stage. Possible reasons for this result are the noun's imagery and/or concreteness effect (Weiland et al., 2014; Forgacs et al., 2015; Schmidt-Snoek et al., 2015). As N400 is associated with simulation and mental imaging of perceptual motion (Holcomb et al., 1999), semantic concreteness influences the N400 effect. Concrete nouns tend to elicit larger N400s than abstract nouns (West and Holcomb, 2000; Kanske and Kotz, 2007; Adorni and Proverbio, 2012; Barber et al., 2013). In the present study, only the object of the literal-concrete sentence is a concrete noun, which may lead to a stronger N400 effect in the frontal and central brain areas induced by the object of the literal-concrete sentence compared to the other three types of sentences.

CONCLUSION

Based on the debate that whether the processing of verbal metaphor is relied entirely on perceptual motion simulation or involves abstract sense processing systems, the present study was designed to assess the cognitive mechanism of Chinese verbal metaphor processing, which is rarely investigated. Firstly, the processing mechanism of the metaphorical meaning of the verbal metaphor is a neural pattern that combines literal-concrete meaning and literal-abstract one, involving more cognitive resources than processing the above two meanings, respectively. It supports the simulation-abstraction hybrid view and thus has expanded the existing theoretical accounts of metaphor. Secondly, the processing of verbal metaphors is a dynamic process with gradual change. The concrete and abstract meanings are activated instantly according to the unfolding of sentence meaning at different stages, facilitating the extraction and integration of the metaphorical meaning of the expression. Therefore, the present findings can further clarify the processing process of the Chinese verb metaphor. Furthermore, the current research believes that when it comes to the mental representation of semantic embodiment, the processing of verbal metaphors should have cross-cultural consistency. The perceptual motion is also involved in the processing of Chinese verbal metaphors. However, when it comes to specific language characteristics, English language expressions are mostly abstract, while Chinese languages are characterized by image expressions. The differences of linguistic conventions may lead to the different performance of sensorimotor and abstract semantic involvement in the processing of Chinese verbal metaphors compared with English or other Indo-European languages characters. Specifically, although verbs in the verbal metaphorical sentences are concrete verbs, the activation of abstract characteristics is more dominant in English expressions (Tian et al., 2020), while in Chinese verbal metaphors, whether literal-concrete meaning will be processed prior has not been studied. The current study can enlighten whether verbal metaphors comprehension have cross-cultural consistency between different languages, and whether the perceptual motion system has a unique role different from that in the processing of Chinese verbal metaphors.

In addition, the findings of the present study on the neural processing mechanism of Chinese verbal metaphors in SOV sentences can provide some pedagogical implications for Chinese language teaching practices and learning. This study shows that the sensory-motor system is involved in the processing of Chinese metaphorical comprehension. Thus, the embodiment of language learning should be emphasized in Chinese teaching and learning. That is, language teachers and learners need to pay much attention to the use of multimodal resources, especially a variety of body movements, such as gesture and body posture, or embodied simulation of literal meanings. The use of these resources can enhance non-literal language learning and memory. In the process of Chinese metaphorical comprehension, besides the sensory-motion simulation of concrete concepts, the abstract sense stored in long-term memory will also be activated to promote the immediate integration of metaphorical meanings. Therefore, language learners should pay attention to the accumulation of knowledge in daily life to acquire figurative meanings more quickly.

However, there are limitations in the study. First, a limited amount of context is available before the target verb. Future studies using more extensive context (e.g., a complete sentence or multiple sentences) can potentially provide valuable insights into the effects of context and activation of meanings. Second, despite the high temporal resolution of event-related potentials (ERPs) technology used in the current study, its spatial resolution is not as accurate as other neuroimaging technologies, making it difficult to determine the specific brain regions activated during the comprehension of verbal metaphor. Therefore, subsequent studies need to further explore the neural mechanism of verbal metaphor processing with the help of other neuroimaging technology, to provide more direct evidence of brain activation for sensorimotor simulation. Third, this study did not distinguish or compare the embodied dimensions of verbs (e.g., verbs related to mouth, leg, hand, etc.), and the embodied dimensions involved were not comprehensive enough. Whether there would be differences in verb metaphors of different types of verbs could be further discussed in future studies. As such, future studies should consider syntax context, verb type and the variety of experimental methods.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YL: conceptualization, methodology, investigation, formal analysis, writing - review & editing; XL: investigation, formal analysis, writing - original draft; YiW: writing - review & editing; HW: drafting the work and revising it critically for important intellectual content. YuW: conceptualization, funding acquisition, supervision, writing - review & editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Project of Humanities and Social Sciences from Ministry of Education in China (20YJC190023) and Outstanding Young Scientific Research Team Project in humanities and social sciences of Zhengzhou University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.877997/full#supplementary-material

REFERENCES

Adorni, R., and Proverbio, A. M. (2012). The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia* 50, 880–891. doi: 10.1016/j.neuropsychologia.2012.01.028

Al-Azary, H., and Katz, A. N. (2020). Do metaphorical sharks bite? Simulation and abstraction in metaphor processing. *Mem. Cogn.* 49, 557–570. doi: 10.3758/s13421-020-01109-2

Barber, H. A., Otten, L. J., Kousta, S. T., and Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain Lang.* 125, 47–53. doi: 10.1016/j.bandl.2013.01.005

Bardolph, M., and Coulson, S. (2014). How vertical hand movements impact brain activity elicited by literally and metaphorically related words: an ERP study of embodied metaphor. *Front. Hum. Neurosci.* 8:1031. doi: 10.3389/fnhum.2014.01031

Barsalou, L. W., Simmons, W. K., Barbey, A. K., and Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* 7, 84–91. doi: 10.1016/s1364-6613(02)00029-3

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., et al. (2014). Creating metaphors: The neural basis of figurative language production. *NeuroImage* 90, 99–106. doi: 10.1016/j.neuroimage.2013.12.046

Boulenger, V., Shtyrov, Y., and Pulvermüller, F. (2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *NeuroImage* 59, 3502–3513. doi: 10.1016/j.neuroimage.2011.11.011

Cameron, L., and Gibbs, R. (2008). "Metaphor and talk," in *The Cambridge handbook of metaphor and thought*. ed. R. W. Gibbs (United Kingdom: Cambridge University Press), 197–211.

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., and Chatterjee, A. (2012). From novel to familiar: tuning the brain for metaphors. *NeuroImage* 59, 3212–3221. doi: 10.1016/j.neuroimage.2011.11.079

Chatterjee, A. (2010). Disembodying cognition. *Lang. Cogn.* 2, 79–116. doi: 10.1515/LANGCOG.2010.004

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., and Seidenberg, M. S. (2011). The neural career of sensory-motor metaphors. *J. Cogn. Neurosci.* 23, 2376–2386. doi: 10.1162/jocn.2010.21596

Desai, R. H., Conant, L. L., Binder, J. R., Park, H., and Seidenberg, M. S. (2013). A piece of the action: modulation of sensory-motor regions by action idioms and metaphors. *NeuroImage* 83, 862–869. doi: 10.1016/j.neuroimage.2013.07.044

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Feng, Y., and Zhou, R. (2021). Does embodiment of verbs influence predicate metaphor processing in a second language? Evidence From Picture Priming. *Front. Psychol.* 12:759175. doi: 10.3389/fpsyg.2021.759175

Forgacs, B., Bardolph, M. D., Amsel, B. D., DeLong, K. A., and Kutas, M. (2015). Metaphors are physical and abstract. ERPs to metaphorically modified nouns resemble ERPs to abstract language. *Front. Hum. Neurosci.* 9:28. doi: 10.3389/fnhum.2015.00028

Gallese, V., and Lakoff, G. (2005). The brain's Concepts: The role of the Sensory-motor system in Conceptual Knowledge. *Cogn. Neuropsychol* 22, 455–479. doi: 10.1080/02643290442000310

Gentner, D., and Asmuth, J. (2019). Metaphoric extension, relational categories, and abstraction. *Lang. Cogn. Neurosci.* 34, 1298–1307. doi: 10.1080/23273798.2017.1410560

Gentner, D., and France, I. (1988). "The verb Mutability effect: Studies of the Combinatorial Semantics of Nouns and Verbs," in *Lexical Ambiguity resolution*. ed. R. W. Gibbs (Netherlands: Elsevier), 343–382.

Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind Lang.* 21, 434–458. doi: 10.1111/j.1468-0017.2006.00285.x

Holcomb, P. J., Kounios, J., Anderson, J. E., and West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 721–742. doi: 10.1037/0278-7393.25.3.721

Jamrozik, A., McQuire, M., Cardillo, E., and Chatterjee, A. (2016). Metaphor: bridging embodiment to abstraction. *Psychon. Bull. Rev.* 23, 1080–1089. doi: 10.3758/s13423-015-0861-0

Ji, H., Qi, S., Xu, S., Chen, J., Dai, D. Y., Li, Y., et al. (2020). The role of animacy in metaphor processing of mandarin Chinese: An event-related potential (ERP) study. *J. Neurol.* 56:100915. doi: 10.1016/j.jneuroling.2020.100915

Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044

Lai, V. T., Howerton, O., and Desai, R. H. (2019). Concrete processing of action metaphors: evidence from ERP. *Brain Res.* 1714, 202–209. doi: 10.1016/j.brainres.2019.03.005

Lauro, L. J. R., Mattavelli, G., Papagno, C., and Tettamanti, M. (2013). She runs, the road runs, my mind runs, bad blood runs between us: literal and figurative motion verbs: An fMRI study. *NeuroImage* 83, 361–371. doi: 10.1016/j.neuroimage.2013.06.050

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004

Obert, A., Gierski, F., and Caillies, S. (2018). He catapulted his words from the dais: An ERP investigation of novel verbal metaphors. *J. Neurol.* 47, 59–70. doi: 10.1016/j.jneuroling.2018.02.008

Qu, F.-B., Yin, R., Zhong, Y., and Ye, H. S. (2013). Motor perception in language comprehension: perspective from embodied cognition. *Adv. Psychol. Sci.* 20, 834–842. doi: 10.3724/sp.J.1042.2012.00834

Raposo, A., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47, 388–396. doi: 10.1016/j.neuropsychologia.2008.09.017

Richardson, D. C., Spivey, M. J., Barsalou, L. W., and McRae, K. J. (2003). Spatial representations activated during real-time comprehension of verbs. *Cogn. Sci.* 27, 767–780. doi: 10.1016/S0364-0213(03)00064-8

Rutter, B., Kroger, S., Hill, H., Windmann, S., Hermann, C., et al. (2012). Can clouds dance? Part 2: an ERP investigation of passive conceptual expansion. *Brain Cogn.* 80, 301–310. doi: 10.1016/j.bandc.2012.08.003

Schmidt-Snoek, G. L., Drew, A. R., Barie, E. C., and Agauas, S. J. (2015). Auditory and motion metaphors have different scalp distributions: an ERP study. *Front. Hum. Neurosci.* 9:126. doi: 10.3389/fnhum.2015.00126

Semino, E., Steen, G., and Gibbs, R. W. Jr. (2008). "Metaphor in literature," in *The Cambridge handbook of metaphor and thought*. ed. R. W. Gibbs (United Kingdom: Cambridge University Press), 232–246.

Shen, Z. Y., Tsai, Y. T., and Lee, C. L. (2015). Joint influence of metaphor familiarity and mental imagery ability on action metaphor comprehension: An event-related potential study. *Lang. Linguist.* 16, 615–637. doi: 10.1177/1606822x15583241

Tian, L., Chen, H., Zhao, W., Wu, J., Zhang, Q., de, A., et al. (2020). The role of motor system in action-related language comprehension in L1 and L2: An fMRI study. *Brain Lang.* 201:104714. doi: 10.1016/j.bandl.2019.104714

Weiland, H., Bambini, V., and Schumacher, P. B. (2014). The role of literal meaning in figurative language comprehension: evidence from masked priming ERP. *Front. Hum. Neurosci.* 8:583. doi: 10.3389/fnhum.2014.00583

West, W. C., and Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *J. Cogn. Neurosci.* 12, 1024–1037. doi: 10.1162/08989290051137558

Wilson, N. L., and Gibbs, R. (2007). Real and Imagined body Movement Primes Metaphor Comprehension. *Cogn. Sci.* 31, 721–731. doi: 10.1080/15326900701399962

Yang, G., Bradley, K., Huq, M., Wu, D.-L., and Krawczyk, D. (2013). Contextual effects on conceptual blending in metaphors: An event-related potential study. *J. Neurol.* 26, 312–326. doi: 10.1016/j.jneuroling.2012.10.004

Yin, R., Su, D., and Ye, H. (2013). Conceptual metaphor theory: basing on theories of embodied cognition. *Adv. Psychol. Sci.* 21, 220–234. doi: 10.3724/sp.J.1042.2013.00220

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 14 July 2022
doi: 10.3389/fpsyg.2022.915952

Check for updates

Facilitative Effects of Embodied English Instruction in Chinese Children

Connie Qun Guan[1,2]* *and Wanjin Meng*[3]*

[1] School of Foreign Studies, Beijing Language and Culture University, Beijing, China, [2] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States, [3] Department of Moral, Psychological and Special Education, China National Institute of Education Sciences, Beijing, China

OPEN ACCESS

Edited by:
Naomi Sweller,
Macquarie University, Australia

Reviewed by:
Claudia Repetto,
Catholic University of the Sacred Heart, Italy
Airil Haimi Mohd Adnan,
MARA University of Technology, Malaysia

***Correspondence:**
Connie Qun Guan
conniequnguan@blcu.edu.cn
Wanjin Meng
1205747017@qq.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *08 April 2022*
Accepted: *16 June 2022*
Published: *14 July 2022*

Citation:
Guan CQ and Meng W (2022) Facilitative Effects of Embodied English Instruction in Chinese Children. Front. Psychol. 13:915952. doi: 10.3389/fpsyg.2022.915952

Research into the lexical quality of word representations suggests that building a strong sound, form, and meaning association is a crucial first step for vocabulary learning. For children who are learning a second language (L2), explicit instruction on word morphology is generally more focused on whole word, rather than sub-lexical, meaning. Though morphological training is emphasized in first language (L1) vocabulary instruction, it is unknown whether this training facilitates L2 word learning through sub-lexical support. To test this, we designed three experimental learning conditions investigating embodied morphological instruction [i.e., hand writing roots (HR), dragging roots (DR), gesturing roots (GR)] to compare against a control condition. One hundred students were randomly assigned to the four experimental groups. Pre- and post-tests examining knowledge of word meanings, forms, and sounds were administered. Results of mixed linear modeling revealed that three embodied morphological instruction on roots enhanced L2 vocabulary learning. Hand writing roots facilitated sound-meaning integration in all category-tasks for accessibility to word form and one task for word sound-form association. By contrast, GR facilitated meaning-based learning integration in two out of three category tasks for word form-meaning association. Chunking and DR facilitated meaning-based integration in one out of three category tasks for word form-meaning association. These results provide evidence that the underlying embodied morphological training mechanism contributes to L2 vocabulary learning during direct instruction. Future directions and implications are discussed.

Keywords: embodied cognition, language instruction, explicit morphological training, English as a second language, handwriting, gesture

INTRODUCTION

"Learning by doing" has long been recognized as an effective learning mechanism since the early 19th century (Dewey and Authentic, 1938). It refers to learning from experiences resulting directly from one's own actions (Reese, 2011). Learning to read through bodily engagement with the written form of language could be facilitatory for vocabulary acquisition for children (James, 2010; Kontra et al., 2012; Guan et al., 2021). Embodied learning is defined as the connection between bodily movements and cognitive abilities. Embodied learning shows the relation among movements and academic achievement; in other words, it is the relation of mind and body (Foglia and Wilson, 2013). Embodied language learning emphasizes the usage of motor or sensory movements to process and understand language material, verbs, nouns, or sentences (Buccino and Mezzadri, 2015). Word learning, especially action verbs, often co-occur with bodily movements or visual sensations, contributing to strengthening the link between sensorimotor programs and linguistic concepts (Vukovic and Shtyrov, 2014). The embodied experience might play a significant causal role in second language processing as well, in which a strong bond between context, sensory-motor experience and language was established (Pulvermüller and Fadiga, 2010).

Meaning-based instruction emphasizes the association among lexical constituents centered on the semantic representation of a word. However, contemporary teaching paradigms tend to focus more on phonological approaches such as repetition rather than paying explicit attention to lexical forms (De Jong et al., 2015). High-quality lexical representation requires efficient word recognition and strong association among lexical constituents such as form, sound, and meaning (Ehri, 2014; Stafura and Perfetti, 2014). In the native language acquisition process, mapping high-quality lexical and sub-lexical representations lays the foundation for subsequent acquisition of new words (Perfetti and Harris, 2013). Strong associations among these lexical constituents contribute to literacy in the first language (L1; Perfetti and Harris) and in a second language (L2; Gunderson and D'Silva, 2016). A critical, unresolved question is whether and to what extent explicit embodied instruction could lead to the sound-form-meaning association of high-quality lexical representations involving hand/body-related movements such as hand writing, chunking and dragging, and gesture. In the current study, we examine the question of how children learning English as a Second Language (ESL) can best use bodily engagement to enrich mental representations of the forms, meanings, and sounds of new vocabulary.

The goal of the current study is to provide empirical evidence for explicit embodied instruction in word learning in an L2 context. Presently, there is a wide variety of ways to transfer bodily engagement into learning (for a related discussion, see Skulmowski and Rey, 2018). A large part of embodied learning research is concerned with instructional techniques involving learners' entire bodies (e.g., Johnson-Glenberg et al., 2014; Lindgren et al., 2016). However, other studies have focused on the potential uses of alternative embodied phenomena in educational contexts, including (1) gesturing (Goldin-Meadow, 2011; Pouw et al., 2014), (2) hand writing (De Koning and Tabbers, 2013; Wakefield and James, 2015), and (3) physically manipulating target objects during reading (Glenberg et al., 2013).

Facilitating Letter and Word Recognition *via* Hand Writing

Hand writing facilitates visual word recognition and influences symbol learning by creating a network that includes both sensory and motor brain systems (Guan et al., 2011, 2015). Compared to non-motor practice, handwriting training produces faster learning and greater generalization to untrained tasks than previously reported. Handwriting practice leads to learning of both motor and symbolic letters (Wiley and Rapp, 2021). For example, hand writing Chinese characters improves Chinese second language learners' ability to write Chinese characters and understand their spatial structure (Hsiao et al., 2015). Hand writing facilitates understanding of symbols by virtue of its environmental output, supporting the notion of developmental change through brain-body-environment interactions (Li and James, 2016). Neuroscience research has shown that brain mechanisms supporting visual letter categorization respond more strongly to letters after hand writing of those letters (Kersey and James, 2013). Handwritten letters can facilitate early letter comprehension due to the ability of handwriting to improve visual-motor coordination (Zemlock et al., 2018). Taken together, these results suggest that handwriting practice plays a key role in the brain networks underlying letter perception and lexical learning. However, previous studies have mostly explored the effects of hand writing in comparison with typing, viewing, etc. (Li and James, 2016; Zemlock et al., 2018), with few studies investigating whether hand writing associated with morphological structures at the sub-lexical level facilitate whole word learning.

Facilitating Word Recognition *via* Chunking and Dragging

A *chunk* is defined as a sequence with the property that elements within the sequence, but not elements outside the sequence, predict each other (Cohen and Adams, 2001). Chunking refers to the organization and presentation of information into easily identifiable and manageable groups or units so that people can efficiently and effectively understand and process a message (Lorenz and Tizón-Couto, 2019). In cognitive psychology, chunking is a process by which individual pieces of a set of information are broken down and then combined into a meaningful whole. Information is grouped in large chunks to increase the short-term retention of material, thereby bypassing the limited capacity of working memory (Thalmann et al., 2019). Chunking, as a memory strategy, facilitates retention, word segmentation, reading, information processing, language acquisition, and comprehension (Gobet et al., 2001; Lorenz and Tizón-Couto, 2019). For example, reading skills depend in part on the development of higher-order perceptual units, sometimes referred to as chunking. This concept was proposed to help explain the so-called word comprehension effect, or the idea that more letters can be understood at once if they form a word than if they are unrelated (Stennett et al., 1973). Studies of Chinese character chunking have shown that while certain cortical areas of the dorsal and ventral pathways are activated during chunking, activation of early and higher visual cortical areas is inconsistent (Luo et al., 2006). Using fMRI to investigate the relationship between different frequencies and the chunk status of derived words (e.g., government, worthless), it was found that relative frequency affects the early stages of processing, thus supporting the notion of chunking based on the frequency of use elicited (Blumenthal-Dramé et al., 2017).

Several studies also support the idea that performing tracing activities with fingers and other simple hand movements can aid learning processes and language development (Freeman et al., 2011; Brooks and Goldin-Meadow, 2016). Researchers have measured hand movements on a screen to understand the dynamics of a broad range of psychological processes. Hand-tracking can provide unusually high-fidelity, real-time motor traces of the mind (Freeman et al., 2011). Spivey et al. (2005) asked participants to move the computer mouse from the bottom-center of the screen to the top-left or top-right corners, and hand movements showed a continuous attraction toward the distractor before settling into the correct alternative. Dragging as a kind of hand movement can facilitate word learning by dragging the individual chunks of letters or parts into words. From the perspective of motor skill development, one study tested participants on their use of click-drag-click or drag-and-drop motions and results showed that drag-and-drop task was faster and more accurate for children (Donker and Reitsma, 2007). Another study investigated whether point-and-click versus drag-and-drop interactions during game play had an effect on achievement and motivation (Inkpen, 2001) and results suggested that point-and-click was quicker, more accurate, and generally easier for children, and had a positive impact on motivation and success. Previous motor experience has been shown to affect how language is understood and processed—playing hockey can enhance one's ability to understand language about hockey, apparently because brain areas normally used to perform an act become highly involved in understanding language about that act (Pulvermüller, 2005; Beilock et al., 2008). This study used physical manipulation (dragging) to help better understand the morphological structure and facilitate word recognition process. Gestalt psychology postulates that the perception of an entire string does not involve strong activation of its components, but that isolated components strongly evoke the whole.

Most studies on chunking and dragging have focused on L1 speakers (Perruchet et al., 2014); studies of chunking in L2 learners are relatively rare, tending to focus on the chunking effects of L1 and L2 segmentation processes (Mauranen, 2009; Franco and Destrebecqz, 2012). To date, very few studies investigate whether chunking and dragging of morphological structures at the sub-lexical level facilitates whole word learning.

Facilitating L1 and L2 Learning *via* Gestures

Gestures are a form of communication in which body movements convey information supplementing information conveyed *via* language. Research on the role of gestures in the early stages of language learning has shown that pointing can be used to facilitate the expression of new ideas and words (Goldin-Meadow and Alibali, 2013). Moreover, gestures facilitate semantic processing of complex narratives in children as well as adults (Austin and Sweller, 2014; Dargue and Sweller, 2018). Gestures are closely associated with concomitant speech processing (Kelly et al., 2015) and are useful in disambiguation (Holle and Gunter, 2007) and communication (Kelly et al., 2015), with gestures conveying phonological information also playing a facilitating role (Holler and Wilkin, 2011). Research has shown that gestures can affect language comprehension in native speakers (Hostetter, 2011). With respect to second language learning, gestures can enhance L2 vocabulary acquisition and retention (Allen, 1995; Tellier, 2008; Kelly et al., 2009; Morett, 2014, 2018) and improve acquisition of novel L2 speech sounds (Morett and Chang, 2014; Zhen et al., 2019; Hoetjes and van Maastricht, 2020; Li et al., 2021). Gestures have been found to influence the three interrelated cognitive processes of communication, encoding, and recall for L2 vocabulary learning (Macedonia, 2014; Morett, 2014). Gestures accompanying L2 speech facilitate its acquisition by contributing to embodied representations, suggesting that gestures should be incorporated into L2 instruction as a learning tool (Macedonia, 2014). Because the majority of previous studies have focused on the phonological and communicative aspects of L2 learning in which gestures play a facilitatory role, little published research to date has investigated whether gestures associated with morphological structures at the sub-lexical level facilitate whole-word learning.

Taxonomy

A large part of embodied learning research is concerned with instructional techniques involving learners' entire bodies (Johnson-Glenberg et al., 2014; Lindgren et al., 2016). However, other studies have focused on the potential uses of alternative embodied phenomena in educational contexts, including (1) gesturing (Goldin-Meadow, 2011; Pouw et al., 2014), (2) hand writing (De Koning and Tabbers, 2013; Wakefield and James, 2015), and (3) physically manipulating target objects during reading (Glenberg et al., 2013).

Skulmowski and Rey (2018) have proposed a more general model based on the two dimensions of integration between bodily engagement and task performance. A 2 (levels of bodily engagement low vs. high) × 2 (levels of meaning association: low vs. high) grid was proposed according to the model/taxonomy to assess corresponding learning outcomes. Specifically, these are classified as follows: (1) hand writing (decoding + form + meaning mapping, and binding); (2) dragging (cue-based manipulation + form embedding); and (3) action (semantic + bodily engagement), as well as other embodied actions ranging from subtle finger typing to a global concept of learning by using the whole body. Their model posits that bodily activities could be integrated into learning tasks and that bodily action should be associated with enriched semantic representations of to-be-learned materials.

According to this taxonomy, the degree of the facilitative effect of instruction increases incrementally as the degree of bodily engagement increases. The current study implemented three types of experimental conditions, namely hand movement, and physical manipulation (i.e., chunking and dragging), and action. These conditions can be categorized as high integrated bodily engagement, low integrated bodily engagement, and low incidental bodily engagement, respectively. These three conditions allow us to compare embodied learning activities ranging from limited to full-body movement systematically and informatively. The activation of motor systems (body engagement or hand movement) facilitates semantic processing. Therefore, this experiment tested whether handwriting, dragging, and acting as embodied manipulative strategies in Latin root morphology training could improve child ESL learners' vocabulary learning.

Theories and Practices of Morphological Training

Current Model of Morphological Processing and Instruction

Morphosyntax is the system of rules that govern how morphemes are combined to form words that express different meanings. Morphemes are the smallest meaningful units in a given language, and include three morphemic patterns: inflectional affixes, derivational affixes, and roots. Morphological awareness is the metalinguistic awareness that words are made up of meaningful units. It is the foundation of generative word processing. Morphological awareness enables the development of effective and intrinsic morphological processing and problem-solving strategies. There is extensive evidence that morphological awareness is of great importance for word learning for both L1 and L2 English learners, as it is associated with literacy outcomes such as decoding (e.g., Singson et al., 2000), word identification (e.g., McCutchen et al., 2009), spelling ability (e.g., McCutchen and Stull, 2015), and reading comprehension (Crosson and McKeown, 2016).

The present study is based on Schreuder and Baayen's (1995) theoretical model of morphological processing, which indicates how morphemes are established as memory representations. According to this theoretical model, form-meaning associations between orthographic strings and their corresponding meanings begin to be established *via* frequent encounters. Once learners discover the redundant form-meaning link, the memory representation (or concept node) is established. When the orthographic pattern is encountered again and again, the strength of the memory representation becomes stronger and stronger. In addition, when an unfamiliar morphologically complex word is encountered, representations of morphological structure can be activated in one's memory. In this sense, this theoretical model demonstrates how morphemes can be used in morphological analysis to infer the meanings of unfamiliar words.

This theoretical model does not adequately account for the use of Latin roots in word learning. In the view of Schreuder and Baayen (1995), frequency and transparency are essential for activating and strengthening memory representations, but Latin roots are not always redundant and transparent enough to activate these representations. For example, some Latin roots are not transparent in semantics, which may affect their accessibility [e.g., The latin root "cal" refers to "to call," but the word "calvary" (ordeal) has nothing to do with the meaning of "to call." So sometimes the latin roots can be non-transparent]. Other Latin roots may have different phonological and orthographic forms, which may affect the accessibility of the morphological constituent. Therefore, the current framework merits further elaboration and extension to establish memory representations of Latin roots. Using morphological information about Latin roots to establish form-meaning associations and infer meanings of unfamiliar words has been overlooked in previous morphology research.

Latin Roots Training

Latin roots hold great value in facilitating L2 vocabulary learning (Crosson and Moore, 2017). Seventy-five percent of academic words in English are Latinate (Lubliner and Hiebert, 2011), with their main semantic components being bound roots. The present study draws upon Crosson and McKeown's (2016) terminology, applying the term "Latin root" to roots, most often from Latin, that are the semantic basis of English words but are not free-standing words in English. We apply the term "root-related words" to freestanding words that share a Latin root (e.g., *fluent*, *flush*, *fluid* and *influence* are all root-related words, as all contain the Latin root *flu*).

Latin roots are often the core unit of a word, such as *bene* in *benefit*. If a learner does not know the meaning of the root *bene* (good or well), knowing the derivational affix *dis* may be useless. Thus, knowledge of Latin roots is essential for morphological analysis. Root-related words share phonemes, graphemes, and semantics and thus demonstrate consistent recurring relations among sound, spelling, and meaning. Identification and segmentation of root-related words may contribute to more effective acquisition of word pronunciation, spelling, and reading comprehension.

Current Study

The current study explored how children learning ESL best learn words through explicit embodied instruction by testing three conditions: writing the words (hand writing roots; HR), manipulating a chunk consisting of letters by dragging it to make up a word (dragging roots; DR), and demonstrating the meanings of words *via* gestures (gesturing roots; GR). The current study aims to address the extent to which these embodied approaches to L2 vocabulary learning contribute to high-quality lexical representations. We examined these three embodied learning conditions in comparison to a control, word-meaning only (WMO) condition. We conducted explicit morphological training to investigate how children learning ESL use morphological information from Latin roots to enhance their vocabulary learning.

The primary research questions of the present study were: (1) Whether or not embodied morphological training enhances the quality of lexical representations for novel words in children learning ESL? (2) How/to what extent embodied morphological training instruction differs in enhancing the quality of lexical representations for novel words in children learning ESL due to different aspects of embodied engagement? In relation to these two questions, we predicted that: (1) morphological training using Latin roots will improve word learning by strengthening semantic and orthographic representations of words; (2) hand writing Latin roots will enhance morphological awareness and establish form-meaning associations, facilitating the use of Latin roots to infer the meanings of unfamiliar words; (3) embodied morphological manipulative strategies will support word learning *via* morphological training with Latin roots.

MATERIALS AND METHODS

Participants

One hundred Chinese-speaking children in grades 3 through 6 learning ESL were recruited for participation in the present study. All participants attended public elementary schools in Ningbo City, Zhejiang Province, Southeastern China and had been studying English since grade 3, when they were an average of 9 years old. Participants were randomly assigned to four training conditions in equal proportions: HR, chunking and DR, GR, and word meaning only (WMO; control condition). There were no significant differences in age between groups ($p > 0.01$). English language proficiency was assessed using the Peabody Picture Vocabulary Test (PPVT) and a standardized reading measure before training. There was no statistically significant difference in language proficiency between the groups as measured by these two assessments ($ps > 0.05$). Questionnaires on the children's Chinese and English learning background (Li et al., 2021) were also completed by their teachers, confirming that students had no learning difficulties in either L1 or L2.

Training Materials

Eighteen Latin roots and 44 words containing those roots were selected from the Corpus of Contemporary American English (COCA, Davies, 2008). The selected roots comprised four independent Latin roots having no variant and seven roots with one variant. The frequency rank for the selected words ranged from 257 (*receive*) to 56715 (*biped*). There were 18 (=4 + 7*2) sets of trained roots and their 44 corresponding words.

The Latin roots tested in the present study were selected mainly based on frequency, family size, and the sub-categories to which they belonged. All 18 Latin roots and 44 words were taught in each of four learning conditions. The following criteria guided the selection of Latin roots: (1) the most frequent morphographs in English according to Becker et al. (1980) corpus of morphographic units in the 26,000 highest frequency words in English (Davies, 2008); (2) having as many root-related words as possible, such as *exclaim* and *proclaim*. In order to collect more root-related words, low-frequency words could also be selected if they were of shorter length (e.g., *biped* contains only five letters, but ranks 56715 in COCA); (3) belonging to one of the following five sub-categories: body action, body part, mind and emotion, substance, or abstract concept.

Local school teachers were also asked to rate the familiarity of all 44 selected words to guarantee that the target words were all unfamiliar to the students, as highly frequent encounters with any morphological roots could impact whether and how quickly that root would be recognized in other words (Nagy and Hiebert, 2011).

Training Conditions and Sessions

All training was conducted in class by four different instructors. The instructors were four researchers who received systematic training in all facets of the project. All four conditions (HR, DR, GR, and WMO) lasted 8 days and covered 18 Latin roots and 60 words. The training materials for all learning conditions followed the design flowchart (see **Figure 1**) and were identical except for the learning activity. While the learning activities in each training condition differed, each target word was presented a total of eight times during the activity and review in each condition. **Figure 1** presents the instruction flowchart and 8 days of learning conditions for each word taught in the four conditions. Each target word was presented six times in the main session when the novel words were introduced through the instructor's presentation of slides and the learning condition (i.e., hand writing, dragging, gesture, or memorizing the meaning). Participants were exposed to each target word two times in the follow-up session, which began by reviewing the words taught in the previous session.

In the *HR group:* Participants were told to write each root by hand in the blank (e.g., write *grad* in the blank of "____uate"). Each root was written eight times in total.

In the *DR condition*: Participants were instructed to place a sticker printed with the root in the corresponding word blank. For example, on Day 1, "*grad*" *and* "*gress*" were printed on four stickers, and participants placed "*grad*" or "*gress*" onto "_____uate," "up_____," "re_____" or "ag____sive." This was repeated eight times for each word.

In the *GR condition*: Participants were presented with a video showing a gesture of the root meaning first and then told to produce the same gesture as the video showed. Videos were made in advance with an experienced native English speaking ESL teacher acting out the 18 roots individually. After playing the video, the instructor acted out each of these gestures again and asked the students to do the same eight times for each word in total. For example, *grad* was gestured as two hands climbing steps, and students copied this gesture eight times.

In the *WMO condition*: Participants visually read the words and memorized the word meaning from the definition eight times.

Training Procedure

As shown in **Figure 1**, two Latin roots and their four corresponding new words were introduced in Day 1. For Day 1, there were four steps of training. In *Step 1*, all instructors in each of the four conditions taught the concept of Latin roots briefly by giving an example and explaining their function. In *Step 2*, instructors introduced the new words by prompting participants to guess the meaning of the target Latin roots of the day and confirmed the correct meaning of the root in both Chinese and English. In *Step 3*, instructors conducted the experimental learning conditions (i.e., HR, DR, GR, WMO) in the training session for each word. In *Step 4*, instructors introduced the words used in context and led participants in reading context-based sentences containing the novel words, completing learning activities twice for each of the novel words. In *Step 5*, instructors consolidated the meaning of the words by reinforcing the same learning activities in each representative condition twice. On Day 2, instructors began by reviewing the words taught on the previous day, asking participants to complete

FIGURE 1 | Eight day training flowchart.

the learning activities twice. Therefore, participants in each of the four learning conditions received a total of eight exposures to the target words.

Measures

Data were collected using both qualitative and quantitative measures. Pre- and post-tests for each of the four conditions were conducted before and after the whole training session, respectively. Participants' classroom teachers and two trained research assistants administrated a computerized lexical decision task (20-min duration) and a paper-and-pencil word knowledge test (60-min duration) in a formal classroom setting. There were two versions in which the order of the 10 assessments were counterbalanced. Participants were assigned randomly to either of the two versions. All 10 assessments were divided into four categories based on the degree of association between lexical constituents.

Category I Tasks: Accessibility to Word Form

Lexical Decision (3 Min)

This was an individualized computerized task. The goal of this task was to assess sensitivity to word forms. The stimuli included 48 pseudowords, 24 taught words, and 24 novel words (similar to words used in the Slash-Control Task) controlled for frequency, part of speech, and semantic neighborhood. Pseudowords had been matched on bigram frequency and word length (Martens and de Jong, 2008). The stimuli were presented in E-Prime, and participants completed the task in a quiet computer room. The task instructions were given in Chinese to make sure all participants understood the task requirements. Participants were required to respond as quickly as possible by pressing the "Yes" button if the word was a real word or the "No" button if it was a non-word. The reliability of this task was a = 0.87 for all participants.

Word Generation Task (15 Min)

This was a group-administered paper-and-pencil test. The goal of this task was to assess participants' ability to generate words containing the target Latin roots. Fourteen Latin roots were provided. Participants were asked to generate as many words using the given root as they could from their knowledge within 14 min. One minute per root was provided to complete the task.

Slash Task for Target Words (1.5 Min)

This was a group-administered paper-and-pencil test aimed at assessing participants' ability to use a slash to correctly separate letter strings into words. Words taught in training were presented without spaces, four in a row, in random order. Participants were asked to separate the letters in each string into four distinct words. For example, the following string might be presented: "surviveevidencesuperiornegate." The correct response after slashing separated the string into the words "survive/evidence/superior/negate."

Slash Task for Control Words (1.5 Min)

This task served as a control baseline for the Slash Task for Target Words and was designed to assess participants' ability to separate letter strings into familiar control words. It was also a group administered paper-and-pencil test. The requirements were the same as the Slash Task for Target Words, except that the words selected were familiar before training.

Category II Tasks: Word Form-Sound Association

Dictation Task (5 Min)

This task was a group-administered paper-and-pencil test. The classroom teacher recited all 44 words taught during the previous training session. Participants listened to the pronunciation and wrote down the words on a piece of paper. A total of 1.5 points were given for each word (1 point for correct spelling of the root, 0.5 point for correct spelling of the other portion of the word), with 36 points as a full mark for this task. If a portion of the root of the word was provided, it was scored based upon the proportion of accuracy out of 1 point for a correct spelling of that root.

Category III Tasks: Word Form-Meaning Association

Word Form-Meaning Matching Task (8 Min)

This was a group-administered paper and pencil test designed to evaluate participants' ability to map a word form onto its meaning. Participants were asked to match the word with its corresponding meaning, which was represented by the word's definition in English. There were three sets within the word-definition matching task, with eight words and their corresponding definitions comprising one set. A total 24 word-definition pairs were assessed. All the words were taught during the earlier training session.

Root Match Task (3 Min)

This was a group-administered paper and pencil test. The goal was to evaluate whether participants had learned the meaning of the Latin roots taught in the training. Participants matched Latin roots with their corresponding meanings. There were three sets within the root-meaning matching task, with six Latin roots and their corresponding meanings comprising one set. A total of 18 roots taught in the training sessions were assessed.

Translation Task (5 Min)

This was a group-administered paper and pencil test. The goal was to assess whether participants could recall the meanings of words taught during the training session. They were required to write the Chinese meaning of each English word taught in training. All 44 words were assessed.

Category IV Tasks: Morphological Awareness Control

Chinese Morphological Awareness Task 1 (12 Min)

This was an individualized task. This morphological compound task (Leong et al., 2008) contained two parts that varied in generating left-headed or right-headed two-character morphological compound words with eight base items each. Participants could choose any six base forms to produce as many "right-headed" two-character words as they could in 6 min. Then, they were asked to choose any six base forms to produce as many "left-headed" two-character words as they could in 6 min. For example, given one of the base forms "马" (ma3, horse), students were required to come up with as many compound words as possible, such as 马上 (literal translation: on the top of the horse = *immediately*), 马路 (literal translation: the path for horse riding = *road*), 马头 (*horse's head*), 马车 (literal translation *cart driven by the horse*), etc. Two research assistants scored freely affixed items according to the base character. Inter-rater reliability was 0.97. Cronbach's alpha internal consistency reliability of all the items for this measure was 0.70.

Chinese Morphological Awareness Task 2 (5 Min)

This was an individualized morphological chain task. Participants were required to provide as many two-character compound words from the left-headed base character as possible in 5 min. For example, after the first compound word "苹果 (literal translation: apple *fruit* = *apple*)" was given, the following morphological chain could be produced by the students "果园 (literal translation: fruit yard = *orchid*)," "园艺 (yard *arts*)," "艺人 (artistic *person*)," "苹果(person's *heart*)," "心眼 (heart *hole*)," etc. Two research assistants scored the freely affixed items according to the base of the character. Inter-rater reliability was 0.98. Cronbach's alpha internal consistency reliability of all the items for this measure was 0.74.

Procedure

Immediately before and after each training session, all participants were assessed using the 10 tasks described above. After the pre-test, participants were randomly assigned to one of four training groups. Independent-samples t-tests showed no statistically significant differences in pretest performance on any measures between the four training groups ($ps > 0.05$). Training was conducted for 8 days, with one 45-min session per day.

Data Analysis

We analyzed our data using linear-mixed effects models (Baayen et al., 2008; Davies et al., 2017), which can simultaneously account for both participant and item level differences. In mixed-effects models, the unit of analysis is the outcome of an individual trial rather than the average across multiple trials. We examined the dependent measure of eight tasks: the accuracy of each of these eight measures, using a generalized mixed-effects model as the log odds (*logit*) of correctly judging a word, log-transformed to reduce positive skew.

We analyzed participants' true responses in a mixed-effect logit models as a function of two fixed effects[1]: test times (pre vs. post tests, "Learning" in the formula), and condition ("Cond" in the formula). All variables were coded with mean-centered contrast to obtain estimates of main effects analogous to those from a repeated measures ANOVA.

$$log(y_{ijk}) = \gamma_{0000} + \gamma_{1000}(\text{Cond}_{ij} - \overline{Cond}) + \gamma_{2000} (\text{Cond}_{ij} - \overline{Cond})^2 + \gamma_{3000}Learning_k + \gamma_{13000} (\text{Cond}_{ij} - \overline{Cond})Learning_k + \gamma_{23000}(\text{Cond}_{ij} - \overline{Cond})^2 Learning_k + \gamma_{4000}MA_{ij} + \gamma_{14000}(\text{Cond}_{ij} - \overline{Cond})MA_{ij} + \gamma_{24000}(\text{Cond}_{ij} - \overline{Cond})^2 MA_{ij} + u_{ij0} + v_{0j0} + w_{00k} + e_{ijk}$$

where u_{ij0} is the random intercept for subject i (independently sampled from a normal distribution of subject effects with mean 0 and variance τ^2_U), v_{0j0} is the random intercept for classroom j (independently sampled from a normal distribution of classroom effects with mean 0 and variance τ^2_V), w_{00k} is the random intercept for item k (independently sampled from a normal distribution of item effects with mean 0 and variance τ^2_W), and e_{ijk} is a random trial-level error term (independently sampled from a normal distribution with mean 0 and variance σ^2_e). The model of logit accuracy for subject i in classroom j responding to item k was the same except that the trial-level error term was omitted and the dependent measure was $log\ (\frac{y_{ijk}}{1-y_{ijk}})$, where y_{ikj} is the probability of subject i in classroom j responding correctly to item k.

The four ordered conditions (HR, handwrite root; GR, gesture root; DR, chunking and dragging root; WMO, word meaning only) were coded using Helmert contrast, which compares each successive condition group to the mean of the other condition group (see **Figure 2** with the mean of rates of correct responses for each condition group).

In all models, we included both participant, classroom, and item (word) random intercepts to account for both participant differences and, critical to the motivation of the analysis, item differences. We adopted a model-based approach to outlier detection by fitting an initial model, eliminating observations with residuals more than three standard de*via*tions from the mean, then refitting each model (Baayen et al., 2008). This procedure identifies observations that are outliers after considering all fixed and random effects of interest. All models were fit in R using package *lme4* (Bates et al., 2015). Fixed effects were tested using the Wald z test for logit models and the Sattherthwaite approximation to the t distribution for Gaussian models (package *lmerTest*; Kuznetsova et al., 2017), all with an a = 0.05 criterion for significance.

[1]With just four classrooms, the analyses do not intend to get a precise estimate of the variance across classrooms in the population, but to include this random intercept to control for differences across classrooms.

RESULTS

The major results of our analysis are presented in the following order. First, the descriptive statistics of all measures in ten assessments between four conditions show the mean differences between conditions (see **Table 1**). The omnibus test results of main effects of learning and condition, and the learning by condition interaction of all tasks were then reported in **Table 2**. Second, based on the preplanned hypotheses about the differences between specific condition. We then reported the only significant results in learning effects (pre vs. post) and condition effects (between each pair of conditions) of the linear mixed modeling analysis in ten assessments between four conditions (see **Table 3**). Third, we summarized the comparisons between conditions in four category tasks in **Tables 4**, **5**. We present the results in three categories: (1) accessibility of word form, (2) sound-form association, and (3) form-meaning association. These results suggest that the lexical quality of word representations formed by participants in each of the three embodied learning conditions as well as the word-meaning control condition differed significantly.

There were no significant differences between each pair of conditions in the pretest ($ps > 1$) (see **Table 1** for descriptive statistics). Therefore, our preplanned comparisons between pairs of conditions were really to focus on the posttest performance between conditions. The omnibus test of ANOVAs results of main effects of learning and conditions, and the interaction effects of learning by condition of all category tasks were then reported in **Table 2**.

Effects of Learning

First, we examined overall learning tests. The positive intercept term indicates that, overall, participants had improved their performance with the averaged odds 1.33 (95% CI: [1.19, 1.50]) in favor of making more accurate responses toward 1 for each item for all eight measures. Participants responded more accurately to the correct responses more frequently, indicating that they had at least gained their performance overall, regardless of learning conditions. Specifically, the odds of responding true for each item gained from 1.82 to 3.51 times for all tasks. As the major concerns of the research target questions are related to the effects of experimental conditions, we then explored the condition effect further.

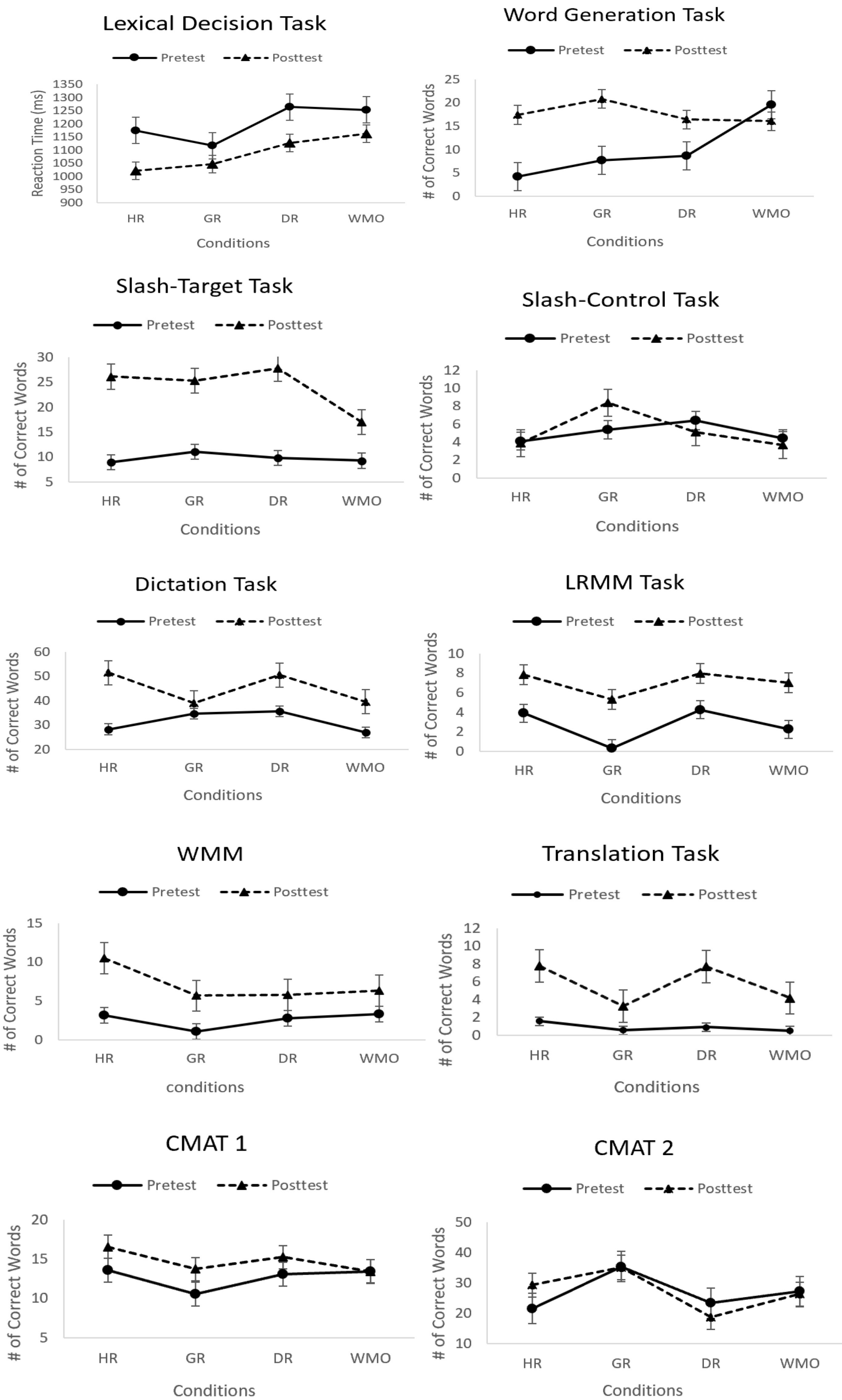

FIGURE 2 | Model-predicted accuracy for eight measures as a function of the partial effects of item-level (pre vs. post) and participant level (four conditions) properties. Error bars depict 95% confidence intervals across subjects. HR, handwrite root; GR, gesture root; DR, dragging root; WMO, word meaning only; LRMM, latin root meaning-matching; WMM, word meaning-matching; CMAT, Chinese morphological awareness task.

TABLE 1 | Descriptive statistics of all measures in 10 assessments between four conditions.

	Conditions							
	HR		GR		DR		WMO	
Tasks	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pretest								
Lexical Decision	1174.78	355.11	1116.52	207.68	1263.50	262.15	1253.17	248.41
Word Generation	4.16	3.10	7.62	5.10	8.60	5.74	19.54	8.92
Slash-Target	8.93	7.58	11.04	11.93	9.76	10.57	9.22	11.26
Slash-Control	4.10	2.43	5.36	7.29	6.40	7.19	4.40	5.00
Dictation	28.21	8.66	34.54	10.23	35.67	8.59	26.98	13.55
LRMM	3.88	2.44	0.28	0.72	4.24	2.89	2.24	2.05
WMM	3.16	1.91	1.05	1.69	2.76	2.54	3.28	2.05
Translation	1.54	2.85	0.52	0.74	0.88	1.09	0.48	0.77
CMAT 1	13.56	4.78	10.52	5.65	13.08	6.66	13.44	5.36
CMAT2	21.44	9.26	35.36	1.78	23.32	10.71	27.08	10.28
Posttest								
Lexical Decision	1021.66	333.18	1045.96	222.25	1127.87	194.08	1162.58	171.02
Word Generation	17.40	12.72	20.81	6.32	16.40	9.51	16.04	8.29
Slash-Target	26.08	22.11	25.26	22.50	27.70	22.07	16.98	16.17
Slash-Control	3.88	3.71	8.36	6.75	5.12	4.50	3.64	3.95
Dictation	51.46	19.78	39.02	13.13	50.49	12.78	39.51	15.63
LRMM	7.84	4.20	5.29	3.19	7.96	5.95	7.00	3.06
WMM	10.48	5.18	5.67	4.82	5.76	4.58	6.32	3.12
Translation	7.76	5.49	3.24	3.35	7.68	6.44	4.16	3.79
CMAT 1	16.56	7.33	13.71	3.07	15.24	5.42	13.40	3.18
CMAT2	29.24	8.66	35.10	5.20	18.64	6.76	26.24	8.62

HR, handwrite root; GR, gesture root; DR, chunking and dragging root; WMO, word meaning only; LRMM, latin root meaning-matching; WMM, word meaning-matching; CMAT, Chinese morphological awareness task.

TABLE 2 | The Omnibus test of ANOVAs results for category tasks.

Category	Measures	Effects	*F*(df)	MSE	*P*	η^2
I Accessibility of Word Form	Lexical decision	Learning	$1.056_{(1,100)}$	100.872	0.893	0.001
		Condition	$0.093_{(3,100)}$	14.918	0.579	0.019
		Learning × Condition	$0.791_{(3,100)}$	76.98	0.429	0.011
	Word Generation	Learning	$44.94_{(1,100)}$	4928.99	<0.001	0.323
		Condition	$5.501_{(3,100)}$	201.132	0.002	0.149
		Learning × Condition	$12.41_{(3,100)}$	1360.86	<0.001	0.284
	Slash Target	Learning	$50.73_{(1,100)}$	21375.1	<0.001	0.351
		Condition	$0.607_{(3,100)}$	100.041	0.612	0.019
		Learning × Condition	$1.78_{(3,100)}$	751.917	0.015	0.054
II	Dictation	Learning	$104.9_{(1,100)}$	16544.59	<0.001	0.528
		Condition	$2.754_{(3,100)}$	360.219	0.047	0.081
		Learning × Condition	$10.70_{(3,100)}$	1686.207	<0.001	0.255
III. Word Form Meaning	Latin-Root Meaning Match	Learning	$73.59_{(1,100)}$	1338.208	<0.001	0.439
		Condition	$10.17_{(3,100)}$	60.214	<0.001	0.245
		Learning × Condition	$0.093_{(3,100)}$	16.918	0.429	0.029
	Word-form Meaning Match	Learning	$66.56_{(1,100)}$	1579.725	<0.001	0.415
		Condition	$8.617_{(3,100)}$	51.421	<0.001	0.216
		Learning × Condition	$3.558_{(3,100)}$	84.440	0.017	0.102
	Translation	Learning	$86.791_{(1,100)}$	1973.325	<0.001	0.480
		Condition	$6.414_{(3,100)}$	47.906	0.001	0.170
		Learning × Condition	$4.161_{(3,100)}$	94.596	0.008	0.117

TABLE 3 | Significant fixed effects estimates from mixed Logit model of correctness with pre-planned comparisons of condition as fixed effects for five tasks.

Task	Effects		Estimate	*SE*	Wald *z*	*p*-value
Lexical decision	Condition Effect	HR vs. WMO	, – 0.60	0.06	– 10.81	< 0.001
		DR vs. WMO	– 0.70	0.06	– 12.62	< 0.001
Word Generation	Condition Effect	HR vs. WMO	0.29	0.11	5.59	< 0.001
		DR vs. WMO	– 0.78	0.11	– 10.81	< 0.001
LRMM	Learning Effect	Pre vs. Post	– 0.52	0.04	– 2.04	0.04
	Condition Effect	GR vs. WMO	– 0.34	0.11	– 2.64	0.01
		GR vs. WMO	– 0.68	0.15	– 4.66	< 0.001
		GR vs. DR	– 0.34	0.15	– 2.24	0.02
WMM	Condition Effect	GR vs. HR	0.29	0.06	5.59	< 0.001
		HR vs. DR	– 0.56	0.06	– 10.11	< 0.001
Translation	Condition Effect	WMO vs. HR	0.12	0.07	1.63	0.10
		GR vs. HR	– 0.09	0.06	– 1.47	0.02
		DR vs. GR	– 0.17	0.07	– 2.46	< 0.001

HR, handwrite root; GR, gesture root; DR, chunking and dragging root; WMO, word meaning only; LRMM, latin root meaning-matching; WMM, word meaning-matching. Only significant effects are presented in the table. All other condition comparisons are not significant, to safe space, so they are not reported in this table.

TABLE 4 | Summary table of simple effect test among conditions.

		Pairwise comparison between conditions					
Tasks	Condition	HR vs. WMO	DR vs. WMO	GR vs. WMO	HR vs. GR	HR vs. DR	GR vs. DR
Pre- vs. posttest							
Lexical Decision	0.046	—	—	—	—	—	—
Word Generation	0.149*	HR > WMO**	DR > WMO*	—	—	—	—
Slash-Target	0.019	—	—	—	—	—	—
Slash-Control	0.049	—	—	—	—	—	—
Dictation	0.080*	—	—	—	—	—	—
LRMM	0.245***	—	—	GR > WMO*	GR > HR***	—	GR > DR***
WMM	0.215*	—	—	—	GR > HR***	HR > DR**	—
Translation	0.17	WMO > HR*	—	—	GR > HR***	—	DR > GR*
CMAT 1	0.025	—	—	—	—	—	—
CMAT 2	0.405***	—	—	—	—	—	—

HR, handwrite root; GR, gesture root; DR, dragging root; WMO, word meaning only; LRMM, latin root meaning-matching; WMM, word meaning-matching; CMAT, Chinese morphological awareness task. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Effects of Condition

What about the effects of different condition? As seen from **Table 4**, learning condition had no main effect on four tasks (e.g., lexical decision, slash-target, slash-control, and dictation), for all other four tasks, we attempted to summarized the comparisons between conditions in the four categories tasks below (see **Table 4**).

The results of the mixed-effects modeling suggested that there was some evidence that participants in the HR condition scored performance better on the tasks overall. For the category tasks of *accessibility to word form*, participants corresponded to a significant 1.25 times (95% CI: [1.02, 1.29]) better in the odds of successfully making correct responses in the HR condition than the GR condition, and a marginal 1.13 times (95% CI: [1.01, 1.27]) increase in the GR condition than the DR condition, and 1.09 (95% CI: [1.00, 1.25]) marginally increase in the DR condition than the WMO condition.

For the category tasks of *word sound-form association*, participants corresponded to a significant 1.15 times (95% CI: [1.02, 1.29]) better in the odds of successfully making correct responses in the HR condition than the WMO condition, and a marginal 1.11 times (95% CI: [1.01, 1.24]) better in the WMO condition than the DR condition, and 1.07 (95% CI: [1.00, 1.21]) marginally better in the DR condition than the GR condition.

TABLE 5 | Summary of comparisons between conditions in four category tasks.

Category	Condition effect
Accessibility to word form	HR > GR > DR > WMO
Word sound-form association	HR > WMO > DR > GR
Word form-meaning association	GR > HR > DR > WMO
Morphological Awareness Control	HR = GR = DR = WMO

For the category tasks of *word form-meaning association*, participants corresponded to a significant 1.35 times (95% CI: [1.04, 1.29]) better in the odds of successfully making correct responses in the GR condition than the HR condition, and a marginal 1.23 times (95% CI: [1.03, 1.26]) better in the HR condition than the DR condition, and 1.03 (95% CI: [1.00, 1.12]) marginally better in the DR condition than the WMO condition. For the Morphological Awareness control task, there were no significant difference between conditions. We summarized the performance below in **Table 5**.

The effect comparison between these four experimental conditions corresponding to four categories of tasks are summarized in **Table 5**. To summarize these results, HR facilitated sound-meaning integration in all category-tasks for accessibility to word form and one task for word sound-form association in comparison with the WMO condition. By contrast, GR facilitated word from-meaning association in two out of three category tasks in comparison with the WMO condition. Chunking and DR facilitated meaning-based integration in one out of three category tasks in comparison with the GR condition, but its facilitative effect in English vocabulary instruction was significantly better in all three category tasks in comparison with WMO condition.

To be concise, it should be noted all three embodied morphological instruction on roots enhanced L2 vocabulary learning. For condition comparison, HR leads to greater effect on sound-form integration, while GR leads to greater effect on form-meaning integration.

DISCUSSION

We took three embodied word learning conditions (HR, DR, GR) and compared them to a WMO condition. The aim was to investigate how children learning ESL use morphological information from Latin roots to improve their L2 vocabulary learning. Our findings were threefold: (1) HR facilitated sound-meaning integration; (2) GR facilitated meaning-based learning integration; (3) DR enhanced meaning-based integration.

Hand Writing Facilitates Sound-Form Meaning Integrated Learning

We found that hand writing facilitates sound-meaning integration. This is compatible with Guan et al.'s (2011) finding that the combination of Chinese handwriting and pinyin together can facilitate learning to read Chinese. Why does HR promote sound-meaning integration? One possible answer is that more listening, speaking, and reading and writing activities in the classroom can strengthen the connection between the phonological forms and meanings of words. In addition, hand writing reflects the fundamental properties of the language system. For skilled learners, activation of the phonological forms of words occurs for all writing systems. Hand writing may lead to the creation of motor representations of spelling patterns that support the development of children's orthographic knowledge. From an applied perspective, the practice of writing words can help children learn to spell (Pritchard et al., 2021). Hand writing is a constructive self-generated learning process or interaction with the construction of sub-lexical letter units of the words being taught. Hand writing engages students in an active decoding process. When learning a decoding system, hand writing is decoded using the language system. In this hand writing process, the visual and auditory systems resonate along with physical or motor embodiment processes to associate phoneme-word correspondences (Guan et al., 2011, 2015; Mizuochi-Endo et al., 2021).

Gesturing Roots Facilitates Meaning-Focused Word Learning

Gesturing roots facilitates the integration of meaning-based learning. This is in line with Singer et al.'s (2008) claim that gestures can help people sensitively capture multiple forms of information representations and sort out the relationships between them and their role in reasoning processes and meaning construction. The reason for this may be that gestures can stimulate intrinsic motivation, in turn improving academic performance (Shakroum et al., 2018). Gestures are part of linguistic output and contain various meanings that are conveyed visually and holistically. Understanding the meaning of gestures is a reasoning process. When the meanings of gestures are understood, it entails construction of the meaning of accompanying language (Parrill and Sweetser, 2004). This also validates Dick et al.'s (2012) finding that children exhibit sensitivity to the meanings of gestures. The developing brain processes the meanings of gestures through different patterns of connectivity in the fronto-temporoparietal network. However, gestural training does not engage decoding practices. In the current study, the learning process in the GR condition did not rely on a phoneme-word correspondence system. Gestures, such as the use of "foot" to refer to the Latin root "ped," did not strengthen the visual and auditory association and their connections to the perceptual-motor system. That is to say, the embodied process of gesturing through pointing the "foot" might not have activated phoneme-word correspondences. Since memory decoding is important for word learning, pointing at "ped" did not access the decoding system. This use of deictic gestures, not the system of language, enters memory. In order to index the meaning of "ped," more work and effort must be expended to create a new "walking" image that maps to the language learning system to facilitate L2 vocabulary acquisition through gesture (Huang et al., 2019).

Dragging Roots Facilitates Form-Associated Learning

Dragging roots enhances form-meaning-sound integration in L2 vocabulary learning for Chinese ESL children. This is consistent with Ellis's (1996) claim that language learners are bound to the orthography and phonology of words. Because it is the word chunks that are dragged, learners need to not only understand the root and word meanings but also use movement to match word components. This chunking and dragging process requires visuomotor integration (Sakai et al., 2003; Bera et al., 2021). Development of visuomotor integration is an important factor in facilitating the development of children's early literacy skills (Dere, 2019). In language acquisition, one of the primary tasks is segmentation of words embedded in a mostly continuous phonological stream (Saffran et al., 1996). Learning a second language, especially with root-based and chunking approaches, requires learners to be adept at inductive reasoning. Learners need to derive general rules from a large number of specific cases (Miao, 2021).

LIMITATIONS

A few limits of the present study should be taken into consideration when interpreting our findings. First, the learning conditions used were limited in duration to only a few days. The number of Latin roots used was only 20 and was extracted from an integrated corpus. Participants were exclusively Chinese elementary school students recruited from a single province. If appropriate, it is also recommended that future studies consider extracting the corresponding high-frequency roots from national vocabulary syllabi for high schools and universities. Further research conducted in more diverse developmental samples across a longer duration is needed to validate the present findings and promote students' second language vocabulary learning.

CONCLUSION

This study aimed to investigate the most effective type of embodied instruction to promote ESL word learning in Chinese children. Four learning methods (HR, GR, DR, WMO) were tested to investigate whether the three training conditions (HR, DR, and GR) embodying word meaning improve the lexical quality of representations of new words compared to the lexical meaning-only control condition (WMO). The differences between these three learning methods were found in those aspects. The study found three main results. One was that handwritten roots promote sound-meaning integration. Second, GR facilitates meaning-based learning integration. Third, DR enhances the form-meaning-sound integration of L2 vocabulary. The study showed that, in the process of teaching vocabulary based on roots, hand writing, gesturing, and dragging should be emphasized so that form-meaning-sound integration can be achieved, contributing to second language vocabulary learning.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Science and Technology, Beijing, China. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CG contributed to the designing, theoretical arguments, data collection, analyses, and writing and editing of the project. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Fundamental Research Funds for the Central Universities at Beijing Language and Culture University (#20YJ020015), Beijing Social Science Key-level Grant (18YYA001), and China National Science Foundations (62077011) awarded to CG.

REFERENCES

Allen, L. Q. (1995). The Effects of Emblematic Gestures on the Development and Access of Mental Representations of French Expressions. *Mod. Lang. J.* 79, 521–529. doi: 10.1111/j.1540-4781.1995.tb05454.x

Austin, E. E., and Sweller, N. (2014). Presentation and production: the role of gesture in spatial communication. *J. Exp. Child Psychol.* 122, 92–103. doi: 10.1016/j.jecp.2013.12.008

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1–48.

Becker, W. C., Dixon, R., and Anderson-Inman, L. (1980). *Morphographic and Root Word Analysis f 26,000 High Frequency Words. Technical Report 1980-1.* Eugene, OR: University of Oregon Follow Through Project.

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., and Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13269–13273. doi: 10.1073/pnas.0803424105

Bera, K., Shukla, A., and Bapi, R. S. (2021). Motor Chunking in Internally Guided Sequencing. *Brain Sci.* 11:292. doi: 10.3390/brainsci11030292

Blumenthal-Dramé, A., Glauche, V., Bormann, T., Weiller, C., Musso, M., and Kortmann, B. (2017). Frequency and chunking in derived words: a parametric fMRI study. *J. Cogn. Neurosci.* 29, 1162–1177. doi: 10.1162/jocn_a_01120

Brooks, N., and Goldin-Meadow, S. (2016). Moving to Learn: how Guiding the Hands Can Set the Stage for Learning. *Cogn. Sci.* 40, 1831–1849. doi: 10.1111/cogs.12292

Buccino, G., and Mezzadri, M. (2015). "Embodied language and the process of language learning and teaching," in *Emotion in Language: Theory–Research–Application*, ed. U. Luedtke (Amsterdam: John Benjamins Publishing Company), 191–208. doi: 10.1075/ceb.10.10buc

Cohen, P., and Adams, N. (2001). "An algorithm for segmenting categorical time series into meaningful episodes," in *International symposium on intelligent data analysis*, eds F. Hoffmann, D. J. Hand, N. Adams, D. Fisher, and G. Guimaraes (Berlin: Springer), 198–207. doi: 10.1007/3-540-44816-0_20

Crosson, A. C., and McKeown, M. G. (2016). Middle School Learners' Use of Latin Roots to Infer the Meaning of Unfamiliar Words. *Cogn. Instr.* 34, 148–171. doi: 10.1080/07370008.2016.1145121

Crosson, A. C., and Moore, D. (2017). When to Take Up Roots: the Effects of Morphology Instruction for Middle School and High School English Learners. *Read. Psychol.* 38, 262–288. doi: 10.1080/02702711.2016.1263699

Dargue, N., and Sweller, N. (2018). Not All Gestures are Created Equal: the Effects of Typical and Atypical Iconic Gestures on Narrative Comprehension. *J. Nonverbal Behav.* 42, 327–345. doi: 10.1007/s10919-018-0278-3

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): One billion words, 1990-2019*. Available Online at https://www.english-corpora.org/coca/ (accessed June 23, 2022).

Davies, R. A., Arnell, R., Birchenough, J. M., Grimmond, D., and Houlson, S. (2017). Reading through the life span: individual differences in psycholinguistic effects. *J. Exper. Psychol. Learn. Memory Cogn.* 43, 1298–1338. doi: 10.1037/xlm0000366

De Jong, N. H., Groenhout, R., Schoonen, R., and Hulstijn, J. H. (2015). Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Appl. Psycholinguist.* 36, 223–243. doi: 10.1017/s0142716413000210

De Koning, B. B., and Tabbers, H. K. (2013). Gestures in instructional animations: a helping hand to understanding non-human movements? *Appl. Cogn. Psychol.* 27, 683–689. doi: 10.1002/acp.2937

Dere, Z. (2019). Analyzing the Early Literacy Skills and Visual Motor Integration Levels of Kindergarten Students. *J. Educ. Learn.* 8, 176–181. doi: 10.5539/jel.v8n2p176

Dewey, J., and Authentic, I. E. L. (1938). *Experiential Learning*. New Jersey: Pentice Hall.

Dick, A. S., Goldin-Meadow, S., Solodkin, A., and Small, S. L. (2012). Gesture in the developing brain. *Dev. Sci.* 15, 165–180. doi: 10.1111/j.1467-7687.2011.01100.x

Donker, A., and Reitsma, P. (2007). Young children's ability to use a computer mouse. *Comput. Educ.* 48, 602–617. doi: 10.1016/j.compedu.2005.05.001

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21. doi: 10.1080/10888438.2013.819356

Ellis, N. C. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *Stud. Second Lang. Acquis.* 18, 91–126. doi: 10.1017/s0272263100014698

Foglia, L., and Wilson, R. A. (2013). Embodied cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 319–325. doi: 10.1002/wcs.1226

Franco, A., and Destrebecqz, A. (2012). Chunking or not chunking? How do we find words in artificial language learning? *Adv. Cogn. Psychol.* 8:144. doi: 10.5709/acp-0111-3

Freeman, J. B., Dale, R., and Farmer, T. A. (2011). Hand in Motion Reveals Mind in Motion. *Front. Psychol.* 2:59. doi: 10.3389/fpsyg.2011.00059

Glenberg, A. M., Witt, J. K., and Metcalfe, J. (2013). From the Revolution to Embodiment. *Perspect. Psychol. Sci.* 8, 573–585. doi: 10.1177/1745691613498098

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5, 236–243. doi: 10.1016/s1364-6613(00)01662-4

Goldin-Meadow, S. (2011). Learning through gesture. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 595–607. doi: 10.1002/wcs.132

Goldin-Meadow, S., and Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annu. Rev. Psychol.* 64, 257–283. doi: 10.1146/annurev-psych-113011-143802

Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F. F., and Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *J. Educ. Psychol.* 103, 509–522. doi: 10.1037/a0023730

Guan, C. Q., Perfetti, C. A., and Meng, W. (2015). Writing quality predicts Chinese learning. *Read. Writ.* 28, 763–795. doi: 10.1007/s11145-015-9549-0

Guan, C. Q., Smolen, E. R., Meng, W., and Booth, J. R. (2021). Effect of handwriting on visual word recognition in chinese bilingual children and adults. *Front. Psychol.* 12:628160. doi: 10.3389/fpsyg.2021.628160

Gunderson, L., and D'Silva, R. A. (2016). "Second language literacy," in *Handbook of Research in Second Language Teaching and Learning*, ed. E. Hinkel (Oxfordshire, UK: Routledge), 490–505.

Hoetjes, M., and van Maastricht, L. (2020). Using Gesture to Facilitate L2 Phoneme Acquisition: the Importance of Gesture and Phoneme Complexity. *Front. Psychol.* 11:575032. doi: 10.3389/fpsyg.2020.575032

Holle, H., and Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *J. Cogn. Neurosci.* 19, 1175–1192. doi: 10.1162/jocn.2007.19.7.1175

Holler, J., and Wilkin, K. (2011). Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue. *J. Nonverbal Behav.* 35, 133–153. doi: 10.1007/s10919-011-0105-6

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychol. Bull.* 137, 297–315. doi: 10.1037/a0022128

Hsiao, H. S., Chang, C. S., Chen, C. J., Wu, C. H., and Lin, C. Y. (2015). The influence of Chinese character handwriting diagnosis and remedial instruction system on learners of Chinese as a foreign language. *Comput. Assist. Lang. Learn.* 28, 306–324. doi: 10.1080/09588221.2013.818562

Huang, X., Kim, N., and Christianson, K. (2019). Gesture and vocabulary learning in a second language. *Lang. Learn.* 69, 177–197. doi: 10.1111/lang.12326

Inkpen, K. M. (2001). Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Trans. Comput.–Hum. Interact.* 8, 1–33. doi: 10.1145/371127.371146

James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Dev. Sci.* 13, 279–288. doi: 10.1111/j.1467-7687.2009.00883.x

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., and Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: two science studies. *J. Educ. Psychol.* 106, 86–104. doi: 10.1037/a0034008

Kelly, S. D., McDevitt, T., and Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Lang. Cogn. Process.* 24, 313–334. doi: 10.1080/01690960802365567

Kelly, S., Healey, M., Özyürek, A., and Holler, J. (2015). The processing of speech, gesture, and action during language comprehension. *Psychon. Bull. Rev.* 22, 517–523. doi: 10.3758/s13423-014-0681-7

Kersey, A. J., and James, K. H. (2013). Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. *Front. Psychol.* 4:567. doi: 10.3389/fpsyg.2013.00567

Kontra, C., Goldin-Meadow, S., and Beilock, S. L. (2012). Embodied learning across the life span. *Top. Cogn. Sci.* 4, 731–739. doi: 10.1111/j.1756-8765.2012.01221.x

Kuznetsova, A., Brockhoff, P., and Christensen, H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26.

Leong, C. K., Tse, S. K., Loh, K. Y., and Hau, K. T. (2008). Text comprehension in Chinese children: relative contribution of verbal working memory, pseudoword reading, rapid automatized naming, and onset-rime phonological segmentation. *J. Educ. Psychol.* 100, 135–149. doi: 10.1037/0022-0663.100.1.135

Li, J. X., and James, K. H. (2016). Handwriting generates variable visual output to facilitate symbol learning. *J. Exp. Psychol. Gen.* 145:298. doi: 10.1037/xge0000134

Li, P., Xi, X., Baills, F., and Prieto, P. (2021). Training non-native aspirated plosives with hand gestures: learners' gesture performance matters. *Lang. Cogn. Neurosci.* 36, 1313–1328. doi: 10.1080/23273798.2021.1937663

Lindgren, R., Tscholl, M., Wang, S., and Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Comput. Educ.* 95, 174–187. doi: 10.1016/j.compedu.2016.01.001

Lorenz, D., and Tizón-Couto, D. (2019). Chunking or predicting–frequency information and reduction in the perception of multi-word sequences. *Cogn. Linguist.* 30, 751–784. doi: 10.1515/cog-2017-0138

Lubliner, S., and Hiebert, E. H. (2011). An Analysis of English–Spanish Cognates as a Source of General Academic Language. *Biling. Res. J.* 34, 76–93. doi: 10.1080/15235882.2011.568589

Luo, J., Niki, K., and Knoblich, G. (2006). Perceptual contributions to problem solving: chunk decomposition of Chinese characters. *Brain Res. Bull.* 70, 430–443. doi: 10.1016/j.brainresbull.2006.07.005

Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Front. Psychol.* 5:1467. doi: 10.3389/fpsyg.2014.01467

Martens, V. E. G., and de Jong, P. F. (2008). Effects of repeated reading on the length effect in word and pseudoword reading. *J. Res. Read.* 31, 40–54. doi: 10.1111/j.1467-9817.2007.00360.x

Mauranen, A. (2009). Chunking in ELF: expressions for managing interaction. *Intercult. Pragmat.* 6, 217–233. doi: 10.1515/iprg.2009.012

McCutchen, D., and Stull, S. (2015). Morphological awareness and children's writing: accuracy, error, and invention. *Read. Writ.* 28, 271–289. doi: 10.1007/s11145-014-9524-1

McCutchen, D., Logan, B., and Biangardi-Orpe, U. (2009). Making Meaning: children's Sensitivity to Morphological Information During Word Reading. *Read. Res. Q.* 44, 360–376. doi: 10.1598/rrq.44.4.4

Miao, C. (2021). The Analysis and Research on the Root-Based and Chunking Approach—Based on the Theory of Multiple Intelligences. *Front. Educ. Res.* 4, 100–108. doi: 10.25236/FER.2021.040719

Mizuochi-Endo, T., Itou, K., Makuuchi, M., Kato, B., Ikeda, K., and Nakamura, K. (2021). Graphomotor memory in Exner's area enhances word learning in the blind. *Commun. Biol.* 4:443. doi: 10.1038/s42003-021-01971-z

Morett, L. M. (2014). When Hands Speak Louder Than Words: the Role of Gesture in the Communication, Encoding, and Recall of Words in a Novel Second Language. *Mod. Lang. J.* 98, 834–853. doi: 10.1111/modl.12125

Morett, L. M. (2018). In hand and in mind: effects of gesture production and viewing on second language word learning. *Appl. Psycholinguist.* 39, 355–381. doi: 10.1017/s0142716417000388

Morett, L. M., and Chang, L.-Y. (2014). Emphasising sound and meaning: pitch gestures enhance Mandarin lexical tone acquisition. *Lang. Cogn. Neurosci.* 30, 347–353. doi: 10.1080/23273798.2014.923105

Nagy, W. E., and Hiebert, E. H. (2011). "Toward a Theory of Word Selection," in *Handbook of Reading Research, Volume IV*, eds M. L. Kamil, P. D. Pearson, E. B. Moje, and P. P. Afflerbach (Milton Park: Routledge), 414–430. doi: 10.4324/9780203840412.ch17

Parrill, F., and Sweetser, E. (2004). What we mean by meaning: conceptual integration in gesture analysis and transcription. *Gesture* 4, 197–219. doi: 10.1075/gest.4.2.05par

Perfetti, C. A., and Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Lang. Learn. Dev.* 9, 296–316. doi: 10.1080/15475441.2013.813828

Perruchet, P., Poulin-Charronnat, B., Tillmann, B., and Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta Psychol.* 149, 1–8. doi: 10.1016/j.actpsy.2014.01.015

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., and Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Front. Psychol.* 5:359. doi: 10.3389/fpsyg.2014.00359

Pritchard, V. E., Malone, S. A., and Hulme, C. (2021). Early handwriting ability predicts the growth of children's spelling, but not reading, skills. *Sci. Stud. Read.* 25, 304–318. doi: 10.1080/10888438.2020.1778705

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582. doi: 10.1038/nrn1706

Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360. doi: 10.1038/nrn2811

Reese, H. W. (2011). The learning-by-doing principle. *Behav. Dev. Bull.* 17, 1–19. doi: 10.1037/h0100597

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996). Word segmentation: the role of distributional cues. *J. Mem. Lang.* 35, 606–621. doi: 10.1006/jmla.1996.0032

Sakai, K., Kitaguchi, K., and Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Exp. Brain Res*. 152, 229–242.

Schreuder, R., and Baayen, R. H. (1995). "Modeling morphological processing," in *Morphological Aspects of Language Processing*, ed. L. B. Feldman (Mahwah, NJ: Lawrence Erlbaum Associates, Inc), 257–294. doi: 10.1002/0470018860.s00254

Shakroum, M., Wong, K. W., and Fung, C. C. (2018). The influence of gesture-based learning system (GBLS) on learning outcomes. *Comput. Educ.* 117, 75–101. doi: 10.1016/j.compedu.2017.10.002

Singer, M., Radinsky, J., and Goldman, S. R. (2008). The role of gesture in meaning construction. *Discourse Process.* 45, 365–386. doi: 10.1080/01638530802145601

Singson, M., Mahony, D., and Mann, V. (2000). The relation between reading ability and morphological skills: evidence from derivational suffixes. *Read. Writ.* 12, 219–252. doi: 10.1023/a:1008196330239

Skulmowski, A., and Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cogn. Res. Princ. Implic.* 3:6. doi: 10.1186/s41235-018-0092-9

Spivey, M. J., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10393–10398. doi: 10.1073/pnas.0503903102

Stafura, J. Z., and Perfetti, C. A. (2014). Word-to-text integration: message level and lexical level influences in ERPs. *Neuropsychologia* 64, 41–53. doi: 10.1016/j.neuropsychologia.2014.09.012

Stennett, R. G., Smythe, P. C., and Hardy, M. (1973). Visual perception of word chunks and beginning reading. *Can. J. Behav. Sci.* 5, 280–289. doi: 10.1037/h0082353

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture* 8, 219–235. doi: 10.1075/gest.8.2.06tel

Thalmann, M., Souza, A. S., and Oberauer, K. (2019). How does chunking help working memory? *J. Exp. Psychol. Learn. Mem. Cogn.* 45:37. doi: 10.1037/xlm0000578

Vukovic, N., and Shtyrov, Y. (2014). Cortical motor systems are involved in second-language comprehension: evidence from rapid mu-rhythm desynchronisation. *NeuroImage* 102, 695–703. doi: 10.1016/j.neuroimage.2014.08.039

Wakefield, E. M., and James, K. H. (2015). Effects of learning with gesture on children's understanding of a new language concept. *Dev. Psychol.* 51, 1105–1114. doi: 10.1037/a0039471

Wiley, R. W., and Rapp, B. (2021). The effects of handwriting experience on literacy learning. *Psychol. Sci.* 32, 1086–1103. doi: 10.1177/0956797621993111

Zemlock, D., Vinci-Booher, S., and James, K. H. (2018). Visual–motor symbol production facilitates letter recognition in young children. *Read. Writ.* 31, 1255–1271. doi: 10.1007/s11145-018-9831-z

Zhen, A., Van Hedger, S., Heald, S., Goldin-Meadow, S., and Tian, X. (2019). Manual directional gestures facilitate cross-modal perceptual learning. *Cognition* 187, 178–187. doi: 10.1016/j.cognition.2019.03.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 22 July 2022
DOI 10.3389/fpsyg.2022.913521

Check for updates

OPEN ACCESS

EDITED BY
Wanjin Meng,
National Institute For Education Sciences, China

REVIEWED BY
Gaoyan Zhang,
Tianjin University,
China
M. Van Hulle,
KU Leuven, Belgium

*CORRESPONDENCE
Hongjun Chen
chenhj@dlut.edu.cn

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 05 April 2022
ACCEPTED 30 June 2022
PUBLISHED 22 July 2022

CITATION
Sun L, Chen H, Zhang C, Cong F, Li X and Hämäläinen T (2022) Decoding brain activities of literary metaphor comprehension: An event-related potential and EEG spectral analysis.
Front. Psychol. 13:913521.
doi: 10.3389/fpsyg.2022.913521

Decoding brain activities of literary metaphor comprehension: An event-related potential and EEG spectral analysis

Lina Sun[1,2], Hongjun Chen[1]*, Chi Zhang[3], Fengyu Cong[3], Xueyan Li[1] and Timo Hämäläinen[2]

[1]School of Foreign Languages, Dalian University of Technology, Dalian, China, [2]Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland, [3]School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China

Novel metaphors in literary texts (hereinafter referred to as literary metaphors) seem to be more creative and open-ended in meaning than metaphors in non-literary texts (non-literary metaphors). However, some disagreement still exists on how literary metaphors differ from non-literary metaphors. Therefore, this study explored the neural mechanisms of literary metaphors extracted from modern Chinese poetry by using the methods of Event-Related Potentials (ERPs) and Event-Related Spectral Perturbations (ERSPs), as compared with non-literary conventional metaphors and literal expressions outside literary texts. Forty-eight subjects were recruited to make the semantic relatedness judgment after reading the prime-target pairs in three linguistic conditions. According to the ERPs results, the earliest differences were presented during the time window of P200 component (170–260ms) in the frontal and central areas, with the amplitude of P200 for literary metaphors more positive than the other two conditions, reflecting the early allocation of attention and the early conscious experience of the experimental stimuli. Meanwhile, significant differences were presented during the time window of N400 effect (430–530ms), with the waveform of literary metaphors more negative than others in the frontal and central topography of scalp distributions, suggesting more efforts in retrieving conceptual knowledge for literary metaphors. The ERSPs analysis revealed that the frequency bands of delta and theta were both involved in the cognitive process of literary metaphor comprehension, with delta band distributed in the frontal and central scalp and theta band in parietal and occipital electrodes. Increases in the two power bands during different time windows provided extra evidences that the processing of literary metaphors required more attention and effort than non-literary metaphors and literal expressions in the semantic related tasks, suggesting that the cognitive process of literary metaphors was distinguished by different EEG spectral patterns.

KEYWORDS

literary metaphor, event-related potentials, N400, P200, neural oscillations

Introduction

It is broadly agreed that metaphorical expressions in literary texts (hereinafter referred to as literary metaphors) are more novel, creative and richer in meaning compared with those outside literary texts (non-literary metaphors; Katz et al., 1988; Lakoff and Turner, 1989; Goatly, 1997; Semino and Steen, 2008). These researchers postulate that metaphorical expressions in literature, e.g., poetry, are used to extend our understanding of ordinary linguistic resources and bring fresh insights into human knowledge, but no consensus has been reached on how metaphor in literary texts is different from that in other communicative texts.

On the one hand, some previous studies (Nowottny, 1965; Leech, 1969; Tsur, 1992; Short, 1996) proposed the discontinuity between metaphor in and outside literature by addressing highly creative, original, and complex literary examples. They explored the uses of metaphor in specific genres, texts, or authors to demonstrate the effects of a particular linguistic choice in its original context. From this point of view, the distinctiveness of a particular use of literary metaphor is highlighted, while non-literary metaphors are considered as derivatives and less worthy of investigation (Semino and Steen, 2008). On the other hand, Lakoff and Johnson (1980) supported the view of continuity between literary and non-literary metaphors in light of Conceptual Metaphor Theory (CMT), which proposed that metaphor is not merely an adornment or entertaining device in human language but a linguistic and cognitive tool which reflects how an abstract and conceptual domain is cognitively structured. They conceived metaphor in everyday language as primary and metaphor in literature as the creative elaboration of ordinary, non-literary metaphor. CMT has resulted in the re-assessment of the role of metaphor in non-literary context and brought new insights into literary metaphor. Lakoff and Turner (1989) also posited that the metaphorical expressions created by poets were novel uses of conventional conceptual metaphors or everyday metaphorical expressions. They argued that these poets extended our way of thinking and expressions by applying creatively the same metaphorical tools to everyday language.

Metaphor processing models

What is more, some psycholinguistic models were proposed to illuminate the neural mechanisms of metaphor comprehension. One of the frequently cited models is the Graded Salience Hypothesis (GSH; Giora, 1997), positing that it is the degree of salience instead of figurativeness that determines the precedence of access. Meanwhile, saliency is determined by the conventionality, frequency, familiarity and prototypicality of the words, phrases or sentences. The literal meaning of novel metaphors is accessed first because the figurative meaning is less salient than literal ones. In contrast, the figurative meaning of conventional metaphors, which is more salient than literal ones, is encoded before the literal meaning. Thus, as opposed to the traditional theories like Standard Pragmatic View (Grice, 1975; Searle, 1979) which attributed temporal priority to the literal meaning, the GSH conceived that the processing differences were not based on the distinction of literalness or figurativeness, but on the degree of salience (Giora, 1997, 2003; Giora and Fein, 1999). The Career of Metaphor model (Gentner and Bowdle, 2001; Bowdle and Gentner, 2005), offered a unified theoretical framework which illustrated whether metaphors were processed directly depended on the degree of conventionality and linguistic form. This model postulates that the comprehension process for conventional and novel metaphors are different. Novel metaphors are understood as comparisons. There is a shift from comparison with categorization in processing as metaphors become increasingly conventionalized (Gentner and Bowdle, 2001; Bowdle and Gentner, 2005; Arzouan et al., 2007; Lai et al., 2009).

Event-related potentials and event-related spectral perturbations of metaphor processing

In recent decades, many researchers have explored the differences between the comprehension of the metaphorical and the literal expressions (Pynte et al., 1996; Arzouan et al., 2007; Mashal et al., 2007; De Grauwe et al., 2010; Diaz et al., 2011; Bambini et al., 2016). Focusing on literary metaphors, some research interests (Katz et al., 1988; Steen, 1994; Goatly, 1997; Reid and Katz, 2022) have been turned to the cognitive aspects of metaphor comprehension. For instance, Steen (1994) pointed out that literary metaphors differed from journalistic metaphors by measuring various dimensions in English and Dutch. Goatly (1997) found that English poetic metaphors were more novel than metaphorical expressions from other texts. In contrast, in Katz et al.'s (1988) study, two sets of literary and non-literary metaphors were analyzed on ten psychological dimensions, such as the degree of metaphoricity, comprehensibility, and the ease of interpretation, but no substantial differences were presented between the two types of metaphors. In spite of this, only a few studies (Arzouan et al., 2007; Rutter et al., 2012; Chen et al., 2016; Tang et al., 2017; Bambini et al., 2019) have attempted to explore literary metaphors through empirical methods.

Event-Related Brain Potentials, with prominently high temporal resolution, is often used to explore the time course of cognitive mechanisms in metaphor processing. The stimulus-locked ERP component of N400, a negative going component peaking around 400 ms, has been well-studied in recent years. It has been shown that the amplitude of N400 varies systematically with the processing of semantic information. The N400 component was also seen as an index of the ease or difficulty of retrieving stored conceptual knowledge related to a word (Kutas and Federmeier, 2000). Most metaphor studies claimed a higher amplitude of N400 for novel metaphors than conventional metaphors and literal expressions (Arzouan et al., 2007; Lai et al., 2009; De Grauwe et al., 2010; Obert et al., 2018). Meanwhile, many studies reported longer reaction time but lower accuracy in the semantic judgment tasks for novel metaphors than literal expressions (Arzouan et al., 2007; Coulson and Van Petten, 2007; Lai et al., 2009; De Grauwe et al., 2010; Bambini et al., 2019). Moreover, the graded N400 waveforms suggested that the difficulty of metaphor comprehension was associated with the complexity of conceptual mapping and information integration. Besides, some other studies also indicated that conventional metaphors and literal expressions elicited similar amplitudes of N400 (Iakimova et al., 2005; Arzouan et al., 2007) due to their high salience and familiarity. Although the N400 amplitude has been discussed a lot in figurative language studies, most results focused on the analysis of time domain and few studies (Ma et al., 2016; Li et al., 2022) have been reported on the domain of time-frequency through ERSPs.

The visual P200 component, peaking between 150 ms and 275 ms, is a positive-going potential, reflecting the early stages in lexical perception. P200 component is suggested to be correlated with contextual information, like sentence-level constraints or congruity related to target words (Coulson and Brang, 2010), the cognitive processes such as working memory (Lefebvre et al., 2005) and memory processing (Dunn et al., 1998). Some recent studies indicated that P200 component was associated with the early processes of high-level language comprehension, such as humor and irony (Regel et al., 2010; Li et al., 2022). For instance, Regel et al. (2010) showed that the P200 component was influenced by the contextual information of speakers' characteristics in literal and ironic language processing. Some studies about metaphor comprehension (Landi and Perfetti, 2007; Schneider et al., 2014) proposed that P200 component was relevant to the ease of decision making in the meaningful judgment tasks. Others reported that P200 was closely associated with pictograph languages like Chinese characters instead of alphabetic languages (Xie et al., 2016). While the N400 component has been discussed extensively, the components preceding N400 have been rarely reported (Freunberger et al., 2007; Lai et al., 2019). Therefore, the component of P200 is well worth investigating in figurative language comprehension for its specific roles in the language perception network.

Based on traditional ERP studies, many studies have shown that neural oscillations perform an essential role in the modulation and generation of ERPs (Makeig et al., 2002; Gruber et al., 2004; Freunberger et al., 2007; Bastiaansen et al., 2008). For instance, Freunberger et al. (2007) evidenced that theta oscillation reflected the top-down regulating processes of memory and was partly involved in the modulation of P200 component. Bastiaansen et al. (2008) discovered that theta power was crucial in the retrieval of lexical semantic information. Compared with conventional ERP studies that focused on the dimension of time, event-related spectral perturbations (ERSPs) could offer a more comprehensive perspective on the analysis of time-frequency dimension, providing a better account of the electrophysiological responses evoked by visual stimuli. Although it is adopted in some language processing studies, the time-frequency dimensional analysis has been scarcely used in metaphorical studies.

Hypotheses

The current study aims to explore the cognitive and neurophysiological underpinnings of literary metaphors by focusing on the variables on the time domain (ERPs components) and time-frequency domain (ERSPs) as electrophysiological responses to visual-evoked neural activations. Firstly, it is hypothesized that there would be significant differences among three language conditions in the early stage, such as the P200 component, of metaphor processing, because P200 was reported to be more significant in processing complex language materials (Zhao et al., 2011). Secondly, we expected a gradient of N400 amplitude in which literary metaphors would elicited the largest waveform, followed by non-literary metaphors and literal expressions. Unlike literal expressions and highly conventional metaphorical expressions, literary metaphors are novel and unfamiliar to the subjects, it would be more challenging to collect information for meaning integration (Tartter et al., 2002; Lai et al., 2009). Finally, we expected to seek more evidence on the relationship between literary metaphors and ERP responses, and the neural oscillations in different frequency bands.

Materials and methods

Subjects

Forty-eight undergraduates and postgraduates from Dalian University of Technology (Liaoning Province, China) were recruited as paid volunteers to participate in the experiment. It has been confirmed that none of the subjects had participated in any of the pretests in this study. All the subjects were right-handed native Chinese speakers, with normal or correct-to-normal vision and no history of neurological/psychiatric disorders or reading disabilities. Written consent form was obtained from all the participants. They were all informed of the instructions and procedures and were asked to minimize body movements, especially from the head, before the experiment. This experiment was approved by the Research Ethics Committee of Dalian University of Technology. Data from six participants were excluded in the statistic analysis due to low number of correct trials ($n = 4$, the ratio of correct trials is 0.34, 0.42, 0.5, and 0.44, respectively) and noisy EEG data ($n = 2$, the ratio of noisy trials is 0.58 and 0.56), leading to a final number of 42 participants (18 male, 24 female) for further analysis. Age ranged from 19 to 25 years old ($M = 22.43$).

Materials

A total of 150 pairs of Chinese phrases with the structure of stimulus 1 (3–12 Chinese characters) to stimulus 2 (2–7 Chinese characters), in the form of prime to target, were selected for the ERP experiment. These stimuli consist of three categories: literary metaphors, non-literary metaphors, and literal expressions, with

TABLE 1 Sample stimuli in the ERP study.

Category	Prime	Target
Literary metaphor	一张金黄的心 A golden heart	九月 September 冬季 Winter
Non-literary metaphor	树木的医生 The doctor of trees	啄木鸟 Woodpecker 害虫 Pest
Literal expression	德国的首都 Capital of Germany	柏林 Berlin 东京 Tokyo

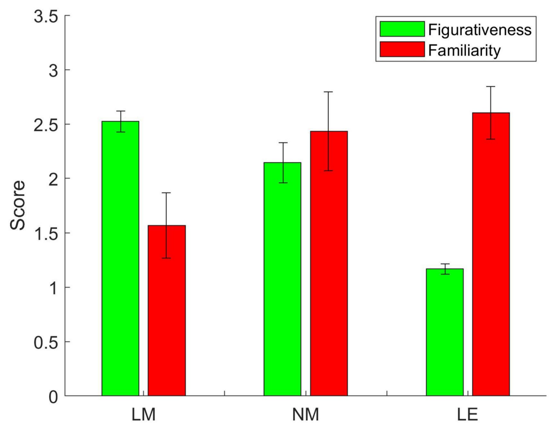

FIGURE 1
Figurativeness and familiarity of the experimental materials (LM, literary metaphors; NM, non-literary metaphors; LE, literal expressions).

50 in each group (see Table 1 for Sample Stimuli). Literary metaphors were natural language extracted from the original context of modern Chinese lyric poems. Comparatively, non-literary metaphors and literal expressions are generally from the news report in Chinese newspapers or magazines. Another group of 50 pairs of phrases which are unrelated in meaning are created as fillers.

This paradigm of prime-target pairs was adopted to examine whether metaphor-comprehension-related neural mechanisms were triggered or not (Sotillo et al., 2004). The prime (stimulus 1) consisted of metaphorical or literal expressions, followed by a target word (stimulus 2) consisting of words or phrases that could or could not be defined by the prime (Sotillo et al., 2004), e.g., "一张金黄的心 (A golden heart)"—"九月 (September)"/"冬季 (Winter)." (The prime-target pairs that are nonrelated in meaning function as fillers). The subjects were required to decide whether Stimulus 1 was accurately described by Stimulus 2 or not. On the one hand, the priming tasks were widely used in figurative language studies, such as allegorical sayings (Zhang et al., 2013) and humor processing (Li et al., 2022). On the other hand, this paradigm was connected with violations of semantic expectation (Kutas and Hillyard, 1980; Barber et al., 2002; Friederici, 2004; Khateb et al., 2010), attention and working memory (Harmony, 2013).

Prior to formal experiment, three pilot surveys were conducted to test the relatedness, figurativeness, and familiarity of the experimental materials. Firstly, 50 raters who did not participate in the formal experiment were enrolled to decide whether the two words or phrases of the prime-target pairs were related with each other in meaning (1 = unrelated, 2 = somewhat related, 3 = highly related). Based on the results, those stimuli rated by at least 90% of the participants as consistent in meaning were selected. After that, another 50 students were asked to judge the figurativeness of the selected stimuli on a 1–3 scale (1 = not figurative, 2 = somewhat figurative, 3 = highly figurative). Expressions with an average of <1.5 were chosen as literal expressions (50 pairs), whereas the expressions with an average of more than 2.5 were selected as literary metaphors and non-literary metaphors (100 pairs). On the final list of stimuli, those selected stimuli were rated by another group of 50 raters on a 1–3 scale (1 = unfamiliar, 2 = somewhat familiar, 3 = highly familiar) regarding their familiarity.

The ANOVA (Analysis of variance) result shows that literary metaphors and non-literary metaphors are much more figurative than literal expressions, and there is a significant difference between the three groups, $F(2, 147) = 1{,}594.04$, $p < 0.01$. According to the result, both non-literary metaphors and literal expressions are more familiar than literary metaphors (Figure 1), and significant differences could be seen clearly between literary metaphors and non-literary metaphors, literary metaphors and literal expressions, $F(2, 147) = 142.79$, $p < 0.01$. Therefore, literary metaphors in lyric poems are more novel compared with those extracted from news report. In contrast, non-literary metaphors are highly conventional according to the result of pilot studies. Although literary metaphors are less familiar, they are still judged as meaningful by the raters, suggesting that literary metaphors are understandable instead of anomalous in meaning.

Experimental procedure

During the whole ERP experiment, the subjects were instructed to sit in a dimly lit sound-attenuated chamber at ~80 cm from a 17-inch computer screen. All the stimuli were presented in white color on a black background. Following the experimental instructions, the subjects were required to read the prime-target pairs silently and judge whether the two words or phrases are semantically related to each other by pressing keys. All the experimental trials were displayed in a pseudo-randomized order to ensure that all the prime-target trials of the same type were not presented consecutively.

To get familiar with the task and procedure, the subjects were instructed with a brief practice of 15 trials of prime-target pairs, which were then not presented in the formal experiment. For each trial, the stimuli were presented in the following time sequence: fixation cross (400 ms), blank (400 ms), prime (2,500 ms), blank (400 ms), target (1,500 ms), and button press (3,000 ms). At the offset of the target word, a 3,000 ms reaction window would be presented. Upon seeing this screen, participants must judge whether these two words or phrases were semantically related (Yes/1, No/3). The inter-trial interval was 1,000 ms before a new trial starts. The overall sequence of events for a trial is illustrated in Figure 2. The formal experiment consisted of 50 trials for each category, with 50 fillers that are unrelated in meaning, leading to a number of 200 trials. The testing session was 30 min with two short breaks of 3 min. The accuracy and response time (RT) in the semantic judgment task were recorded. At last, the subjects with the accuracy rates of lower than 80% were excluded.

EEG recordings and analysis

The EEGs were recorded with an electro cap of 64 Ag/AgCl electrodes according to the 10–20 System of electrode placement. An ANT Neuro EEG amplifier was used to record EEG signals sampled at a digitization rate of 500 Hz. The electrode impedance was kept below 5 kΩ, and the EEG was online referenced to the CPz channel.

In the offline analysis, EEG data were notch filtered at 50 Hz. Next, a digital high-pass filter of 0.5 Hz and a low-pass filter of 30 Hz were applied. After removing the direct current (DC) component, the data were re-referenced to the average of the mastoid references (M1, M2). The ERP epochs from 200 ms before to 1,300 ms after stimulus onset were extracted. Finally, by using the Icasso software (Himberg and Hyvarinen, 2003), independent artifact components (e.g., blinks, movements, etc.) were removed through visual inspection. Data of 6 subjects were excluded due to excessive artifacts.

Event-related potentials

Event-related potentials were analyzed with MATLAB 2019b. First, the individual correct trials whose amplitudes were out of range (max >75 μv, baseline max >30 μv) were rejected, and then the baseline 200 ms before stimulus onset was subtracted from the waveforms. The equal number of trials for each subject under the conditions of LM, NM, and LE (LM = Literary Metaphors; NL = Non-literary Metaphors; LE = Literal Expressions) was adopted based on the minimum number of three condition trials. When the trial number exceeded the minimum number, the trials whose amplitudes were closer to the boundary of the range were removed. Next, trials were averaged across blocks for each subject. The total number of trials across all subjects for each condition was 1,185. The P200 and N400 amplitude and latency were

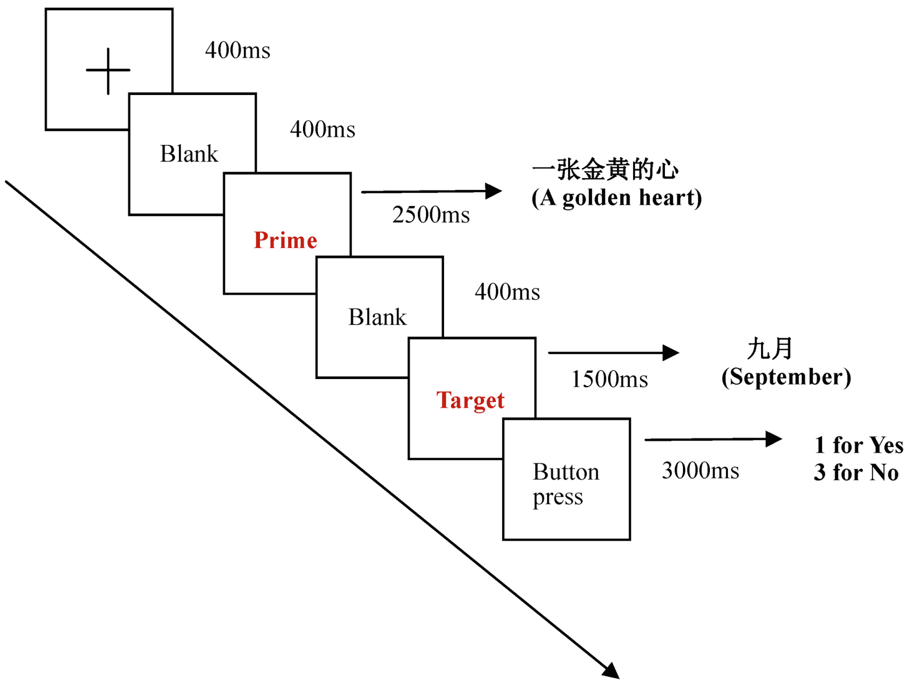

FIGURE 2
Experimental paradigm.

quantified for further analysis. Based on the topographic activations, 15 electrodes (AF7, AF8, F5, F3, F1, Fz, F2, F4, F6, FC1, FCz, FC2, C1, Cz, and C2) were chosen for the N400 analysis. Three electrodes (AF7, AF3, and F5) were chosen for the P200 analysis. The time windows of 170–260 ms and 430–530 ms for the P200 and N400 components were selected. The N400 latency values were calculated as the time of maximum amplitude within the time window of the N400 component (Luck, 2014).

The significance level $p < 0.05$ was used, and all results were reported under the 2-tailed condition. One-way repeated-measures analysis of variance (ANOVA) with three language conditions (LM, NM, and LE) was used to test the hypothesis that LM initiates stronger effect on ERP components such as N400 and P200. Finally, the correlations between performance (accuracy, RTs, and omitted response) and ERP (the amplitude and latency of P200 and N400) were calculated using the Pearson Correlation Coefficient to investigate the association between the behavioral and electrophysiological measures in different language conditions.

EEG spectra

The EEG spectral power was assessed by calculating the event-related spectral perturbation (ERSP) using the continuous wavelet transform (CWT; Guanghui et al., 2018). The complex Morlet wavelet was adopted for the CWT analysis, by which the time-dependent signals were evaluated at each sampling instant with a central frequency band of 1 Hz covering frequencies from 1 to 30 Hz, with a frequency step of 0.1 Hz. Additionally, we normalized the power spectra with the subtraction change from −500 to 0 ms baseline. For quantifying the oscillatory dynamics, we focused on separate time windows in the analysis of two frequency bands. According to the maximum power of the different frequency bands, statistical analysis was performed within the time window of 430–530 ms for the delta band (1–4 Hz) and within the time window of 170–260 ms for the theta band (4–8 Hz). In order to account for the effect of phase-locked (evoked response) activity in the induced oscillations, we also analyzed the induced activations by subtracting the averaged evoked response from each epoch prior to the wavelet analysis.

Results

Behavioral results

Figure 3 illustrates the behavioral performance (accuracy and RT) under different language conditions. Response accuracy was analyzed using one-way ANOVA. There was a significant effect of language conditions ($F(2, 123) = 161.23$, $p < 0.01$, $\eta_p^2 = 0.72$). Multiple comparison tests of Tukey's honestly significant difference (Tukey's HSD) procedure revealed that significant differences were reflected in the comparisons between LM and NM and between LM and LE ($p < 0.01$). Accuracy in the LM condition (mean = 0.6843, SD = 0.1210) was significantly lower than that of the NM condition (mean = 0.9529, SD = 0.0428) and that of the LE condition (mean = 0.9452, SD = 0.0424).

Response Times (RT) to the Probe in the correct trials were analyzed with a one-way ANOVA. There was a significant effect for language conditions ($F(2, 4{,}287) = 39.48$, $p < 0.01$, $\eta_p^2 = 0.02$). Multiple comparison test of Tukey's HSD procedure revealed that significant differences were reflected in the comparisons between LM and NM and between LM and LE ($p < 0.01$). RT in the LM condition (mean = 756.33, SD = 489.92) was significantly longer than that of the NM condition (mean = 639.93, SD = 334.41) and that of the LE condition (mean = 644.06, SD = 349.47).

Event-related potentials results

P200 component

Figure 4A shows the averaged ERP amplitude waveforms with the time window of interest (P200 response at 170–260 ms after stimulus onset), depicted by the gray rectangle. Figure 4C shows the P200 topographies in three language conditions and the P200 topography difference between LM and NM. P200 activity is distributed in the frontal and central areas. The largest differences between LM and NM are mainly situated at the electrodes of AF7, AF3, and F5, where the P200 channels are selected. Figure 4E illustrates the mean values and standard error of the P200 amplitude in the three conditions. The one-way ANOVA reveals significant difference among three language conditions ($F(2, 123) = 3.14$, $p < 0.05$, $\eta_p^2 = 0.07$).

N400 components

Figure 4B shows the averaged ERP amplitude waveforms with the time windows of interest (N400 response at 430–530 ms after stimulus onset), depicted by the gray rectangles. Figure 4D shows the N400 topographies in the three language conditions and the N400 topography difference between LM and NM. N400 activity is distributed in large areas of the forebrain and mid-brain only in LM, which could hardly be detected in the condition of NM and LE. Figure 4F illustrates the mean values and standard error of the N400 amplitude.

The one-way ANOVA reveals a significant effect for language conditions ($F(2, 123) = 6.44$, $p < 0.01$, $\eta_p^2 = 0.09$). Based on the multiple comparison tests, we found the mean N400 amplitude of LM was larger than that of NM and LE. There was no statistically significant N400 response under the conditions of NM and LE.

Event-related spectral perturbation results

Figure 5 illustrates the time-frequency representations (averaged over electrodes AF7, AF8, F5, F3, F1, Fz, F2, F4, F6,

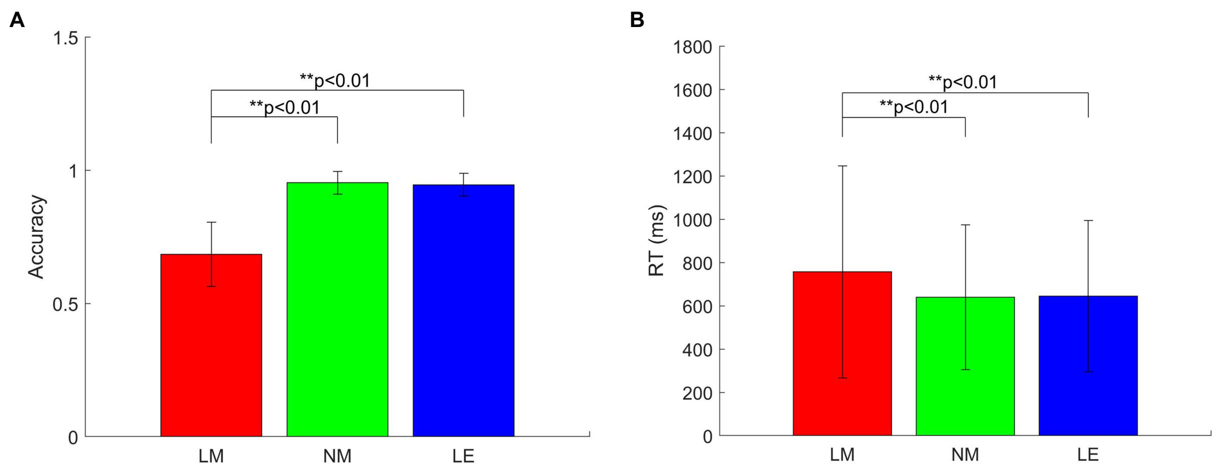

FIGURE 3
Behavioral performance. **(A)** Accuracy; **(B)** reaction time (RT). Averages and standard errors are plotted.

FC1, FCz, FC2, C1, Cz, and C2) in the three experimental conditions. A clear modulation of frequencies of 1–4 Hz is visible in the time window of 430–530 ms. Separable modulations of 4–8 Hz (in the time window of 170–260 ms) appear visually earlier than 1–4 Hz over the three conditions. The oscillations in other frequency bands are not activated with low power values. Therefore, only delta and theta are analyzed in this work. The corresponding frequency bands and time windows are indicated by the dotted-line boxes. Figures 6A, 7A show the power waveforms averaged across the electrodes (referred above) and topographic distribution corresponding to delta band (averaged over 1–4 Hz). The delta oscillations are mainly distributed in the frontal and central areas. Figures 6B, 7B show the power waveforms and topographic distribution of theta band (averaged over 4–8 Hz), with the strongest activations in the parietal and occipital electrodes.

With one-way ANOVA, for delta band power, we found a significant effect for language conditions ($F(2, 123) = 7.51$, $p < 0.01$, $\eta_p^2 = 0.11$) in the frontal and central areas. Based on the multiple comparison tests, we found that delta band power of LM was significantly higher than that of NM and LE. There was no statistically significant difference between the conditions of NM and LE.

For theta band, the significant effect for language conditions ($F(2, 123) = 4.53$, $p < 0.05$, $\eta_p^2 = 0.07$) is only found in the right parietal and occipital electrodes (P8 and PO8). Based on the multiple comparison tests, we found theta band power of LM was significantly higher than that of LE.

Discussion

The main goal of the current study was to draw a clear picture on the cognitive process of novel metaphor comprehension by focusing on the study of literary metaphors from modern Chinese lyric poems. Specifically, this study was designed to evaluate the neural mechanisms of literary metaphors by analyzing the behavioral performances (accuracy and RT), the evoked responses (P200 component and N400 component), and spectral power (delta and theta bands).

Behavioral data

Based on the results of behavioral performance, literary metaphors were proven to be significantly harder to comprehend correctly than non-literary metaphors and literal expressions (see Figure 3), in which the subjects took significantly longer time but achieved lower accuracy for the condition of literary metaphors. In the present study, the experimental stimuli in the group of literary metaphors were evidenced to be less familiar and harder to understand but still evaluated to be related in meaning, leading to the result that subjects spent more time making decisions. The research results were in line with those who reported longer response time and reduced accuracy for novel metaphors in relation to literal sentences (Coulson and Van Petten, 2007; Lai et al., 2009; De Grauwe et al., 2010). Comparatively, the subjects spent nearly the same amount of time in evaluating non-literary metaphors and literal expressions. Meanwhile, the response accuracy of these two language conditions were almost equal to each other. These results demonstrated that the subjects experienced similar comprehension process in approaching these two types of language materials. As a result, based on the behavioral performance, literary metaphors are revealed to be processed through significantly different ways compared with non-literary metaphors and literal expressions, while non-literary metaphors and literal expressions seem to experience similar language comprehension process. However, these results were not

10.3389/fpsyg.2022.913521

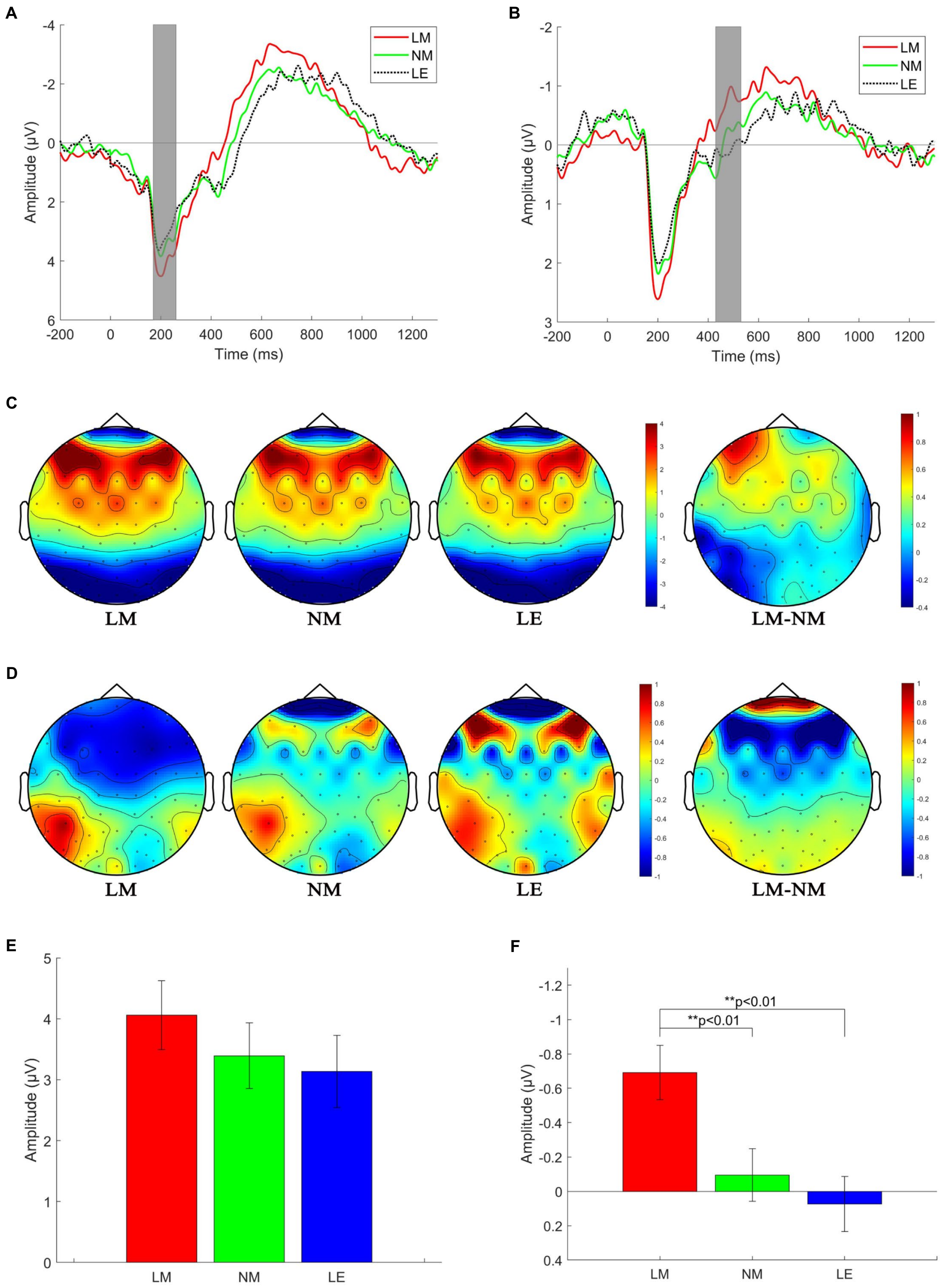

FIGURE 4
ERP responses to three stimulus conditions. **(A)** Grand average ERP of P200 channels. **(B)** Grand average ERP of N400 channels. **(C)** Topographies in P200 time window. **(D)** Topographies in N400 time window. **(E)** Mean values and standard error of the P200 amplitude in the three conditions. **(F)** Mean values and standard error of the N400 amplitude in the three conditions.

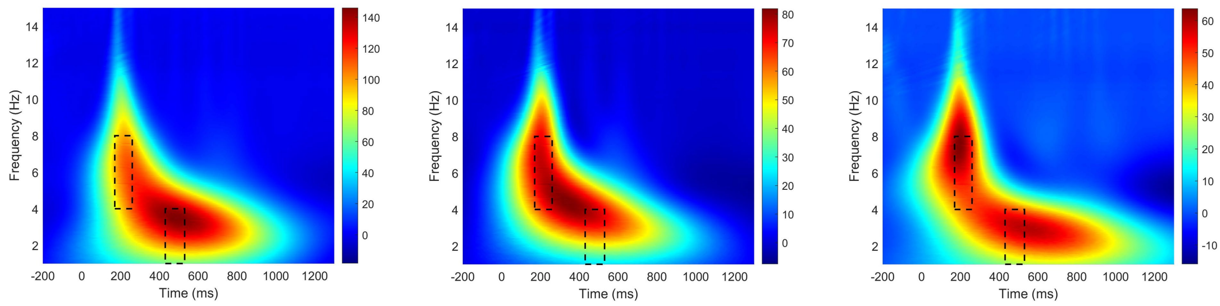

FIGURE 5
Time-frequency representations in the condition of LM, NM, and LE averaged over subjects.

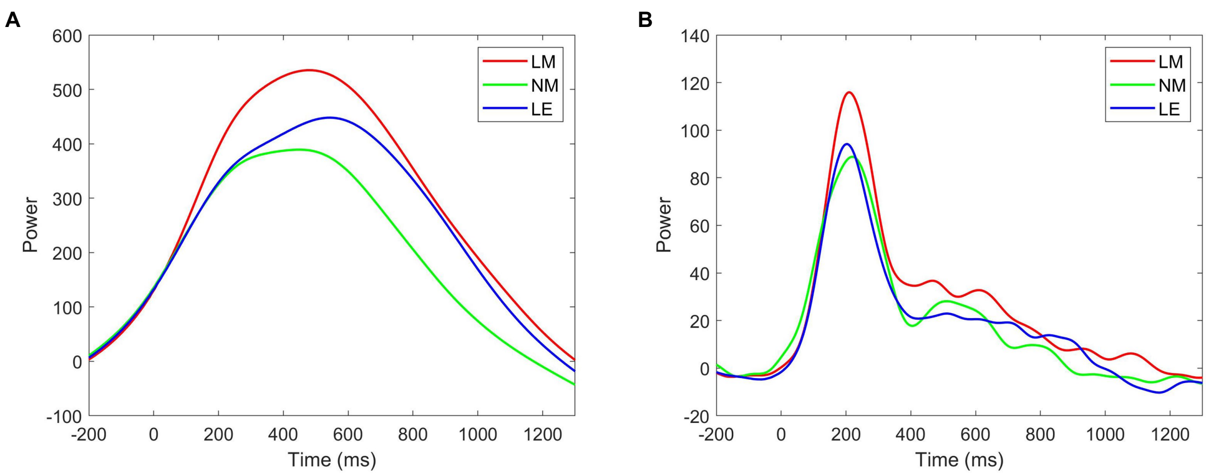

FIGURE 6
Temporal waveforms of power modulation. **(A)** delta; **(B)** theta.

adequate to certify their differences. More information about the neurophysiological underpinnings of metaphor comprehension need to be considered.

Event-related potential data

One of the more significant findings from this study is that the left-frontal P200 component is more positive for literary metaphors than non-literary metaphors and literal expressions. According to McDonough et al. (1992), the P200 component is not only considered to be an exogenous but also an endogenous component, which means that the P200 component may indicate the early sensory stages of item coding such as feature detection (Luck and Hillyard, 1994), selective attention (Hackley et al., 1990) and semantic processing (Federmeier and Kutas, 2002). Our research result about P200 component was compatible with Potts (2004), suggesting that the anterior P200 component was relative to task-relevant stimuli and was especially sensitive to the identification and judgment of experimental stimuli. Similarly, Kim et al. (2008) pointed out that difficult tasks could elicit significantly larger amplitude of P200 (stronger neural activities) than relatively easy tasks. Besides, some researchers (Naatanen and Näätänen, 1992) also reported that the P200 component reflected the early allocation of attention and the early conscious experience of the experimental stimuli. Therefore, the P200 effect was compatible with the assumption that more difficult priming tasks might evoke larger P200 amplitude (Kim et al., 2008; Zhao et al., 2011). In the current study, the stronger P200 effect distributed in the frontal region for literary metaphors might indicate the subjects were attempting to assess the relationship between prime and target words based on the task requirement during the early period of stimulus onset. Specifically, it could be reflected from P200 effect that literary metaphors were more challenging to process, requiring more attention and initial conscious awareness in language processing than non-literary metaphors and literal expressions.

However, this finding is inconsistent with the studies suggesting that P200 amplitude was stronger for the expected context than for the unexpected context (Wlotko and

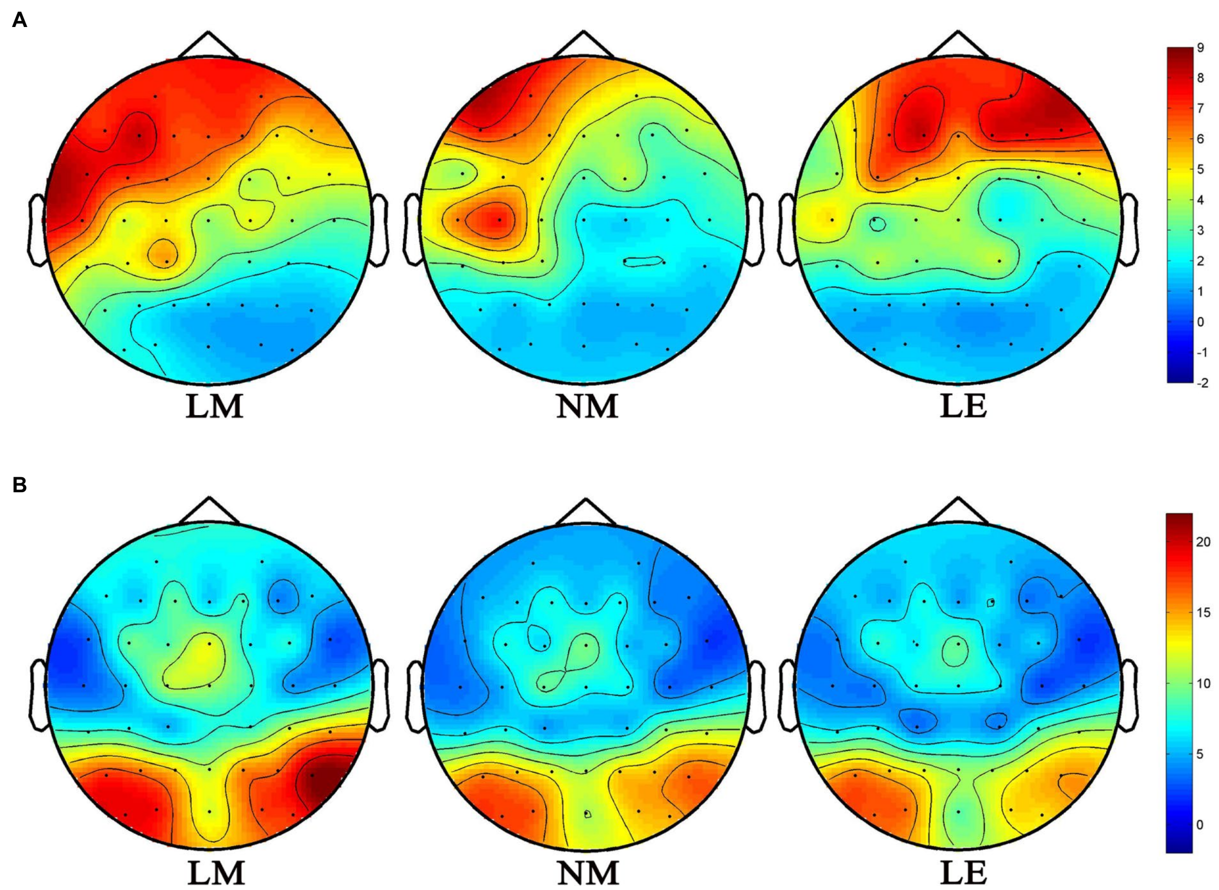

FIGURE 7
Topographies of power modulation. **(A)** delta; **(B)** theta.

Federmeier, 2007). In other words, the stimuli with literal meaning should elicit a stronger P200 component than those with metaphorical sense. This inconsistency may be determined by the experimental paradigm in this study, the prime-target pairs, requiring subjects to make a decision about the relatedness of the two stimuli, which is quite different from the experimental stimuli in previous studies (Federmeier, 2007; Schneider et al., 2014; Li et al., 2022).

The second significant finding is that although there was a graded effect of the three language conditions on the N400 amplitude with literary metaphors eliciting the most negative waveform as reported by previous studies (Lai et al., 2009; Rutter et al., 2012), no significant differences were presented between non-literary metaphors and literal expressions. According to the behavior results, it is apparent from Figure 3 that the differences in accuracy and response time among the three language conditions were significant. The accuracy of literary metaphors is significantly lower, but the reaction time is much longer than the other two groups, in line with the ERP results that the N400 amplitude for literary metaphors is significantly larger than non-literary metaphors and literal expressions. This study suggested that it was more challenging to retrieve conceptual knowledge for literary metaphors than retrieving knowledge of non-literary metaphors or literal expressions in language comprehension. In other words, the construction of metaphorical mappings for literary metaphors was especially complex. The result of N400 responses was consistent with the features of different language materials, in which literary metaphors were always seen as more novel, unexpected and complicated in meaning, resulting in more efforts in establishing mappings between elements within distantly related domains (Arzouan et al., 2007; Goldstein et al., 2012). The interpretation of N400 component was in correspondence with the research finding of Arzouan et al. (2007); Coulson and Van Petten (2007); Lai et al. (2009); De Grauwe et al. (2010) and Bambini et al. (2019) who claimed longer reaction time and decline of accuracy regarding the meaningfulness judgment tasks for novel metaphors than literal expressions.

In terms of topographical distributions, it can be clearly seen that the N400 effect is distributed in the frontal and central area of the scalp for literary metaphors. This region is closely linked to language processing, indicating more efforts in lexical retrieval and semantic integration (Hald et al., 2006). The other two groups, in contrast, have presented few responses of N400. Furthermore, through the topographies of the N400 effect, more regions in the right part of prefrontal cortex were activated, which is reported to be related to figurative language processing (Seger et al., 2000). To this end, as previous neuroimaging studies evidenced (Beeman, 1993; Bottini et al., 1994), the right hemisphere may play a more significant role in connecting distantly related elements in metaphorical language comprehension. The right anterior regions of the N400 effect may signify the semantic integration of literary metaphorical relations. Thus, a further study will concentrate on the right hemisphere advantage in metaphorical language process through source localization algorithms.

According to the topographies during P200 and N400 time windows, significant differences could be observed among the three language conditions. Firstly, for the P200 effect, the topography difference between literary and non-literary metaphors is mainly distributed in the left frontal sites (see Figure 4C, LM-NM), which might play a critical role in the initial stage of meaning integration for literary metaphors. In contrast, no significant differences were found between non-literary metaphors and literal expressions, presenting nearly identical topographies for P200 effect. Secondly, for the N400 effect, the topography difference between literary and non-literary metaphors is mainly distributed in the frontal and central regions (see Figure 4D, LM-NM), which has been reported to be critical in many other studies of novel metaphor comprehension (Bambini et al., 2019; Li et al., 2022). Comparatively, no significant difference was presented between non-literary metaphors and literal expressions, indicating similar processing mechanisms in language comprehension (Iakimova et al., 2005; Arzouan et al., 2007). Meanwhile, our results were compatible with those that treat conventional metaphors as dead metaphors (Goldstein et al., 2012). To sum up, these results evidenced that literary metaphors are significantly different from non-literary metaphors and literal expressions in language comprehension process, while non-literary metaphors and literal expressions presented similar temporal dynamics during comprehension process. To this end, this study was in line with the research findings of Tang et al. (2017) and Bambini et al. (2019), illuminating the brain responses to literary metaphors were similar to those of novel metaphors instead of conventional metaphors.

Event-related spectral perturbations data

The analysis of neural oscillations in EEG signal was proven to be a useful approach, since these data were almost lost in traditional time-locked ERP analysis by averaging single trials. In the present study, the result of ERSPs might help to clarify the neural mechanisms of metaphor comprehension to a greater extent.

Firstly, during the time window of 170–260 ms, literary metaphors were found to evoke significantly stronger increases in the frequency band of theta in comparison with non-literary metaphors and literal expressions in the parietal and occipital area (see Figures 6B, 7B), which is involved in visual cortex (Grill-Spector and Malach, 2004; Guzmán López et al., 2022), in the current study. Although literal expressions were slightly stronger in the frequency band of theta than non-literary metaphors, no significant differences were shown between these two groups (see Figure 6B). Theta increases have been reported to be related to the mental processes such as encoding and memory retrieval (Burgess and Gruzelier, 1997), working memory activation (Gevins et al., 1997; Krause et al., 2000; Deiber et al., 2007) and distribution of attention about target stimuli (Missonnier et al., 2006). Our research result is congruent with previous studies, on the one hand, the experimental design is relevant to visual tasks, as literary metaphors are less familiar and are more difficult to make a judgment, which always require more efforts in visual attention (Missonnier et al., 2006). On the other hand, theta power is associated with working memory process and the power increases with the difficulty of tasks (Weiss et al., 2000; Jensen and Tesche, 2002). Similarly, other researchers (Hagoort et al., 2004; Hald et al., 2006; Davidson and Indefrey, 2007) also claimed that changes of theta activity were connected to violations of semantic expectation, because more efforts were required and more attention were needed during semantic integration. In the present study, literary metaphor was shown to elicit stronger theta power than the other two language conditions, which is consistent with the research finding of Tesche and Karhu (2000), suggesting that theta power was sensitive to the encoding of novel stimuli.

Secondly, during the 430–530 ms time window, a significant effect can be clearly seen among three language conditions in the frequency band of delta, with delta power increases in literary metaphors being significantly larger than the other two conditions in the frontal and central areas of the topographical distributions (see Figures 6A, 7A). According to Fernández et al. (1993), the increase in delta oscillations may be associated with the increasing concentration in semantic evaluation tasks, suggesting that the subjects need to pay more attention to the comprehension of literary metaphors in relation to the other two language conditions. It was also reported that delta activity was consistent with the difficulty of experimental tasks, with more complex task elicit stronger delta power (Harmony et al., 1996). Similarly, as the topographical map shown, delta power mainly distributed in the frontal and central scalp, concurring with the ERP result during the N400 time window that the more difficult tasks (literary metaphors) evoked higher waveform of N400 amplitude. In this regard, our research result is in agreement with previous studies, indicating that delta power was significantly larger for literary metaphors during the time widow of N400, because literary metaphor, which was evidenced to take longer time for the subjects to comprehend, required more attention to deal with the complex activities.

In the current study, the changes in neural oscillations were reflected through different time-frequency topographic maps (see Figure 7), with theta band (4–8 Hz) allocated in the posterior electrodes and delta band (1–4 Hz) distributed in the central and frontal electrodes. Based on previous studies, both delta and theta oscillations functioned in the visual attention tasks (Knyazev, 2012; Keller et al., 2017). As a result, increases in the two power bands for literary metaphors during different time windows evidenced that literary metaphors required more attention and extra effort than non-literary metaphors and literal expressions in the semantic related tasks. In contrast, non-literary metaphors and literal expressions presented similar responses in neural oscillations.

This study tends to provide empirical evidence for the assumption that metaphors in literature are more novel and creative than metaphors outside literature (Semino and Steen, 2008), as readers pay more attention to metaphors in literary texts than to metaphors in non-literary texts (Glicksohn, 1994; Steen, 1994; Goodblatt and Glicksohn, 2002). The research results could be interpreted by the GSH. The salient meaning of non-literary metaphors is the figurative meaning, which could be immediately accessed, as there is no significant difference in lexical access of non-literary metaphors and literal expressions. In contrast, for literary metaphors, the literal meaning is the salient one, more contextual information should be reasonably inferred for the comprehension of the figurative meaning. Our results are also in line with the Career of Metaphor model, on the one hand, the conventionalized figurative meaning are processed through categorization instead of comparison, as there is an existing metaphorical category and no extra efforts were required in the comprehension process. On the other hand, novel metaphors, such as literary metaphors, are processed by establishing correspondences between partially isomorphic conceptual structures of the target and base. After encountering literary metaphors, the initial attempts at categorization failed due to lack of clearly defined category. Therefore, the comparison process begins after discovering that the literal meaning cannot be reasonably applied.

To sum up, this study is consistent with the perspective of continuity between metaphors in and outside literary texts. As Semino and Steen (2008) point out, the metaphorical uses of language need to consider both the uniqueness of specific uses in the language context and how particular uses interact with general conventional patterns, reflecting that the cognitive structures and process might be commonly shared by these two conditions. In other words, literary metaphors often have a conventional basis, and can be seen as extensions and elaborations of conventional metaphors (Semino and Steen, 2008).

Limitations and outlook

Due to the difficulties in measuring electrophysiological responses to literary metaphors during naturalistic comprehension and the complexity of experimental design, the stimuli in this study are words and phrases extracted from the original context. Thus, a future study should be carried out by using more natural materials to examine their temporal and topographical characteristics. For instance, it was proposed that a continuous narrative could be used as experimental stimuli and the neural responses could be recorded along with the discourse. The metaphorical expressions would be time-locked for the analysis of temporal dynamics or frequency power (Bambini et al., 2019). It is a great challenge to move from word level to discourse level to comprehend the differences between literary and non-literary metaphors. Besides, individual differences in comprehending figurative language should be taken into consideration in future studies (Peskin, 2010; Abraham et al., 2021). Previous studies suggested that metaphor comprehension is a long-lasting process, greatly influenced by individual characteristics (Columbus et al., 2015). Accordingly, a further study could focus on the impact of individual differences on figurative language comprehension.

Conclusion

The current study explored the neural mechanisms of literary metaphor comprehension by focusing on the temporal dynamics and neural oscillations of different types of language materials underlying the paradigm of prime-target pairs. Results presented a two-phase language processing procedure, with significantly stronger P200 followed by N400 for literary metaphor condition. Most ERP studies on metaphor comprehension reported N400 and P600 effects, with few studies discussing the role of P200 in metaphor processing. Meanwhile, the neural oscillations of three types of stimuli were in line with the responses of P200 and N400 waveforms. Increases for power bands of delta and theta were found for literary metaphors indicative of statistical differences between metaphors in and outside literature and literal expressions. As for topographical characteristics in ERP study, the frontal and central sites were critical in literary metaphor comprehension. While for the ERSPs results, the delta band and theta band during two different time windows were distributed in significantly different regions, reflecting different roles and functions of the two low-frequency power bands. To sum up, this study reveals a distinctive comprehension process for metaphors in and outside literary texts and literal expressions by uncovering the time courses and EEG spectral patterns. A future study will focus on the study of literary metaphors on discourse level and individual differences.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Research Ethics Committee of Dalian University of Technology. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HC and LS designed the experiment. CZ, LS, and XL analyzed and collected the data. CZ and LS conducted the statistics. LS and CZ wrote the manuscript. HC, FC, and TH revised the manuscript and provided guidance for all the conduction of work. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by "the Fundamental Research Funds for the Central Universities" under Grant DUT20RW401, Grant SIE18RZD2, and the scholarship from China Scholarship Council (No. 202207960001).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.913521/full#supplementary-material

References

Abraham, A., Rutter, B., and Hermann, C. (2021). Conceptual expansion *via* novel metaphor processing: an ERP replication and extension study examining individual differences in creativity. *Brain Lang.* 221:105007. doi: 10.1016/j.bandl.2021.105007

Arzouan, Y., Goldstein, A., and Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Res.* 1160, 69–81. doi: 10.1016/j.brainres.2007.05.034

Bambini, V., Bertini, C., Schaeken, W., Stella, A., and Di Russo, F. (2016). Disentangling metaphor from context: an ERP study. *Front. Psychol.* 7:559. doi: 10.3389/fpsyg.2016.00559

Bambini, V., Canal, P., Resta, D., and Grimaldi, M. (2019). Time course and neurophysiological underpinnings of metaphor in literary context. *Discourse Process.* 56, 77–97. doi: 10.1080/0163853X.2017.1401876

Barber, H., Domınguez, A., and De Vega, M. (2002). Human brain potentials indicate morphological decomposition in visual word recognition. *Neurosci. Lett.* 318, 149–152. doi: 10.1016/S0304-3940(01)02500-9

Bastiaansen, M. C., Oostenveld, R., Jensen, O., and Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain Lang.* 106, 15–28. doi: 10.1016/j.bandl.2007.10.006

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain Lang.* 44, 80–120. doi: 10.1006/brln.1993.1006

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language A positron emission tomography activation study. *Brain* 117, 1241–1253. doi: 10.1093/brain/117.6.1241

Bowdle, B. F., and Gentner, D. (2005). The career of metaphor. *Psychol. Rev.* 112, 193–216. doi: 10.1037/0033-295X.112.1.193

Burgess, A. P., and Gruzelier, J. H. (1997). Short duration synchronization of human theta rhythm during recognition memory. *Neuroreport* 8, 1039–1042. doi: 10.1097/00001756-199703030-00044

Chen, H., Peng, X., Lu, Q., and Wang, H. (2016). ERP differences between literary and non-literary metaphors. *Lang. Cognit. Sci.* 2, 27–53. doi: 10.35534/LCS201602002

Columbus, G., Sheikh, N. A., Côté-Lecaldare, M., Häuser, K., Baum, S. R., and Titone, D. (2015). Individual differences in executive control relate to metaphor processing: an eye movement study of sentence reading. *Front. Hum. Neurosci.* 8:1057. doi: 10.3389/fnhum.2014.01057

Coulson, S., and Brang, D. (2010). Sentence context affects the brain response to masked words. *Brain Lang.* 113, 149–155. doi: 10.1016/j.bandl.2010.02.003

Coulson, S., and Van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension?: ERP evidence from hemifield presentation. *Brain Res.* 1146, 128–145. doi: 10.1016/j.brainres.2007.03.008

Davidson, D. J., and Indefrey, P. (2007). An inverse relation between event-related and time–frequency violation responses in sentence processing. *Brain Res.* 1158, 81–92. doi: 10.1016/j.brainres.2007.04.082

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., and Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia* 48, 1965–1984. doi: 10.1016/j.neuropsychologia.2010.03.017

Deiber, M.-P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., Ibañez, V., et al. (2007). Distinction between perceptual and Attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *J. Cogn. Neurosci.* 19, 158–172. doi: 10.1162/jocn.2007.19.1.158

Diaz, M. T., Barrett, K. T., and Hogstrom, L. J. (2011). The influence of sentence novelty and figurativeness on brain activity. *Neuropsychologia* 49, 320–330. doi: 10.1016/j.neuropsychologia.2010.12.004

Dunn, B. R., Dunn, D. A., Languis, M., and Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain Cogn.* 36, 355–376. doi: 10.1006/brcg.1998.0998

Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505. doi: 10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., and Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia* 40, 730–747. doi: 10.1016/S0028-3932(01)00193-2

Fernández, T., Harmony, T., Rodríguez, M., Reyes, A., Marosi, E., and Bernal, J. (1993). Test-retest reliability of EEG spectral parameters during cognitive tasks: I. Absolute and relative power. *Int. J. Neurosci.* 68, 255–261.

Freunberger, R., Klimesch, W., Doppelmayr, M., and Höller, Y. (2007). Visual P2 component is related to theta phase-locking. *Neurosci. Lett.* 426, 181–186. doi: 10.1016/j.neulet.2007.08.062

Friederici, A. D. (2004). Event-related brain potential studies in language. *Curr. Neurol. Neurosci. Rep.* 4, 466–470. doi: 10.1007/s11910-004-0070-0

Gentner, D., and Bowdle, B. F. (2001). Convention, form, and figurative language processing. *Metaphor. Symb.* 16, 223–247. doi: 10.1080/10926488.2001.9678896

Gevins, A., Smith, M. E., Mcevoy, L., and Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb. Cortex* 7, 374–385. doi: 10.1093/cercor/7.4.374

Giora, R. (1997). Understanding Figurative and literal Language: The Graded Salience Hypothesis. Cognitive Linguistics, 8, 183–206.

Giora, R. (2003). *On Our Mind: Salience, context, and Figurative Language*. New York: Oxford University Press.

Giora, R., and Fein, O. (1999). Irony: context and salience. *Metaphor. Symb.* 14, 241–257. doi: 10.1207/S15327868MS1404_1

Glicksohn, J. (1994). Putting interaction theory to the empirical test: some promising results. *Pragmatics Cognit.* 2, 223–235. doi: 10.1075/pc.2.2.02gli

Goatly, A. (1997). *The Language of Metaphors*. London: Routledge.

Goldstein, A., Arzouan, Y., and Faust, M. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain Lang.* 123, 137–142. doi: 10.1016/j.bandl.2012.09.008

Goodblatt, C., and Glicksohn, J. (2002). Metaphor comprehension as problem solving: an online study of the reading process. *Style* 36, 428–445.

Grice, H. P. (1975). "Logic and conversation," in *Syntax and Semantics 3: Speech Acts*. eds. P. Cole and J. Morgan (New York, NY: Academic Press), 41–58.

Grill-Spector, K., and Malach, R. (2004). The Human Visual Cortex. *Annu. Rev. Neurosci.* 27, 649–677. doi: 10.1146/annurev.neuro.27.070203.144220

Gruber, T., Malinowski, P., and Müller, M. M. (2004). Modulation of oscillatory brain activity and evoked potentials in a repetition priming task in the human EEG. *Eur. J. Neurosci.* 19, 1073–1082. doi: 10.1111/j.0953-816X.2004.03176.x

Guanghui, Z., Tian, L., Chen, H., Li, P., Ristaniemi, T., Wang, H., et al. (2018). Effect of Parametric Variation of Center Frequency and Bandwidth of Morlet Wavelet Transform on Time-Frequency Analysis of Event-Related Potentials.

Guzmán López, J., Hernandez-Pavon, J. C., Lioumis, P., Mäkelä, J. P., and Silvanto, J. (2022). State-dependent TMS effects in the visual cortex after visual adaptation: a combined TMS–EEG study. *Clin. Neurophysiol.* 134, 129–136. doi: 10.1016/j.clinph.2021.08.020

Hackley, S. A., Woldorff, M., and Hillyard, S. A. (1990). Cross-modal selective attention effects on retinal, myogenic, brainstem, and cerebral evoked potentials. *Psychophysiology* 27, 195–208. doi: 10.1111/j.1469-8986.1990.tb00370.x

Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441. doi: 10.1126/science.1095455

Hald, L. A., Bastiaansen, M. C. M., and Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain Lang.* 96, 90–105. doi: 10.1016/j.bandl.2005.06.007

Harmony, T. (2013). The functional significance of delta oscillations in cognitive processing. *Front. Integr. Neurosci.* 7:83. doi: 10.3389/fnint.2013.00083

Harmony, T., Fernández, T., Silva, J., Bernal, J., Díaz-Comas, L., Reyes, A., et al. (1996). EEG delta activity: an indicator of attention to internal processing during performance of mental tasks. *Int. J. Psychophysiol.* 24, 161–171. doi: 10.1016/S0167-8760(96)00053-0

Himberg, J., and Hyvarinen, A. (2003). "Icasso: software for investigating the reliability of ICA estimates by clustering and visualization," in *2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No.03TH8718)*, 259–268.

Iakimova, G., Passerieux, C., Laurent, J. P., and Hardy-Bayle, M. C. (2005). ERPs of metaphoric, literal, and incongruous semantic processing in schizophrenia. *Psychophysiology* 42, 380–390. doi: 10.1111/j.1469-8986.2005.00303.x

Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x

Katz, A. N., Paivio, A., Marschark, M., and Clark, J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor. Symb.* 3, 191–214. doi: 10.1207/s15327868ms0304_1

Keller, A. S., Payne, L., and Sekuler, R. (2017). Characterizing the roles of alpha and theta oscillations in multisensory attention. *Neuropsychologia* 99, 48–63. doi: 10.1016/j.neuropsychologia.2017.02.021

Khateb, A., Pegna, A. J., Landis, T., Mouthon, M. S., and Annoni, J.-M. (2010). On the origin of the N400 effects: an ERP waveform and source localization analysis in three matching tasks. *Brain Topogr.* 23, 311–320. doi: 10.1007/s10548-010-0149-7

Kim, K. H., Kim, J. H., Yoon, J., and Jung, K.-Y. (2008). Influence of task difficulty on the features of event-related potential during visual oddball task. *Neurosci. Lett.* 445, 179–183. doi: 10.1016/j.neulet.2008.09.004

Knyazev, G. G. (2012). EEG delta oscillations as a correlate of basic homeostatic and motivational processes. *Neurosci. Biobehav. Rev.* 36, 677–695. doi: 10.1016/j.neubiorev.2011.10.002

Krause, C. M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M., et al. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clin. Neurophysiol.* 111, 2071–2078. doi: 10.1016/S1388-2457(00)00429-6

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657

Lai, V. T., Curran, T., and Menn, L. (2009). Comprehending conventional and novel metaphors: an ERP study. *Brain Res.* 1284, 145–155. doi: 10.1016/j.brainres.2009.05.088

Lai, V. T., Howerton, O., and Desai, R. H. (2019). Concrete processing of action metaphors: evidence from ERP. *Brain Res.* 1714, 202–209. doi: 10.1016/j.brainres.2019.03.005

Lakoff, G., and Johnson, M. (1980). Conceptual metaphor in everyday language. *J. Philos.* 77, 453–486. doi: 10.2307/2025464

Lakoff, G., and Turner, M. (1989). *More than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago press.

Landi, N., and Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain Lang.* 102, 30–45. doi: 10.1016/j.bandl.2006.11.001

Leech, G. (1969). *A Linguistic Guide to English Poetry. Hong Kong*. London: Longman Group UK Ltd.

Lefebvre, C. D., Marchand, Y., Eskes, G. A., and Connolly, J. F. (2005). Assessment of working memory abilities using an event-related brain potential (ERP)-compatible digit span backward task. *Clin. Neurophysiol.* 116, 1665–1680. doi: 10.1016/j.clinph.2005.03.015

Li, X., Sun, J., Wang, H., Xu, Q., Zhang, G., and Wang, X. (2022). Dynamic impact of intelligence on verbal-humor processing: evidence from ERPs and EROs. *J. Neurolinguistics* 62:101057. doi: 10.1016/j.jneuroling.2022.101057

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique. 2nd ed.* Cambridge, MA: The MIT Press.

Luck, S. J., and Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology* 31, 291–308. doi: 10.1111/j.1469-8986.1994.tb02218.x

Ma, Q., Hu, L., Xiao, C., Bian, J., Jin, J., and Wang, Q. (2016). Neural correlates of multimodal metaphor comprehension: evidence from event-related potentials and time-frequency decompositions. *Int. J. Psychophysiol.* 109, 81–91. doi: 10.1016/j.ijpsycho.2016.09.007

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., et al. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694. doi: 10.1126/science.1066168

Mashal, N., Faust, M., Hendler, T., and Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain Lang.* 100, 115–126. doi: 10.1016/j.bandl.2005.10.005

McDonough, B. E., Warren, C. A., and Don, N. S. (1992). Event-related potentials in a guessing task: The gleam in the eye effect. *Int. J. Neurosci.* 65, 209–219.

Missonnier, P., Deiber, M. P., Gold, G., Millet, P., Gex-Fabry Pun, M., Fazio-Costa, L., et al. (2006). Frontal theta event-related synchronization: comparison of directed attention and working memory load effects. *J. Neural Transm.* 113, 1477–1486. doi: 10.1007/s00702-005-0443-9

Naatanen, R., and Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nowottny, W. (1965). *The Language Poets*. London: Athlone Press, University of London.

Obert, A., Gierski, F., and Caillies, S. (2018). He catapulted his words from the dais: an ERP investigation of novel verbal metaphors. *J. Neurolinguistics* 47, 59–70. doi: 10.1016/j.jneuroling.2018.02.008

Peskin, J. (2010). The development of poetic literacy during the school years. *Discourse Process.* 47, 77–103. doi: 10.1080/01638530902959653

Potts, G. F. (2004). An ERP index of task relevance evaluation of visual stimuli. *Brain and cognition*. 56, 5–13.

Pynte, J., Besson, M., Robichon, F.-H., and Poli, J. (1996). The time-course of metaphor comprehension: an event-related potential study. *Brain Lang.* 55, 293–316. doi: 10.1006/brln.1996.0107

Regel, S., Coulson, S., and Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Res.* 1311, 121–135. doi: 10.1016/j.brainres.2009.10.077

Reid, J. N., and Katz, A. (2022). The RK processor: a program for analysing metaphor and word feature-listing data. *Behav. Res. Methods* 54, 174–195. doi: 10.3758/s13428-021-01564-y

Rutter, B., Kröger, S., Stark, R., Schweckendiek, J., Windmann, S., Hermann, C., et al. (2012). Can clouds dance? Neural correlates of passive conceptual expansion using a metaphor processing task: implications for creative cognition. *Brain Cogn.* 78, 114–122. doi: 10.1016/j.bandc.2011.11.002

Schneider, S., Rapp, A. M., Haeußinger, F. B., Ernst, L. H., Hamm, F., Fallgatter, A. J., et al. (2014). Beyond the N400: complementary access to early neural correlates of novel metaphor comprehension using combined electrophysiological and haemodynamic measurements. *Cortex* 53, 45–59. doi: 10.1016/j.cortex.2014.01.008

Searle, J. R. (1979). "Literal meaning," in *Expression and Meaning: Studies in the Theory of Speech and Irony Processing. Discourse Processes, 35, 241–279. Acts*. ed. J. R. Searle (Cambridge: Cambridge University Press), 117–136.

Seger, C. A., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (2000). Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. *Neuropsychology* 14, 361–369. doi: 10.1037/0894-4105.14.3.361

Semino, E., and Steen, G. (2008). "Metaphor in literature," in *The Cambridge Handbook of Metaphor and Thought*. ed. R. W. Gibbs (Cambridge: Cambridge University Press), *Vol. 6,* 57–70.

Short, M. (1996). *Exploring the Language of Poems, Plays and Drama*. Harlow: Pearson.

Sotillo, M., Carretié, L., Hinojosa, J. A., Tapia, M., Mercado, F., López-Martín, S., et al. (2004). Neural activity associated with metaphor comprehension: spatial analysis. *Neurosci. Lett.* 373, 5–9. doi: 10.1016/j.neulet.2004.09.071

Steen, G. (1994). *Understanding Metaphor in Literature: An Empirical Approach*. London: Longman Publishing Group.

Tang, X., Qi, S., Wang, B., Jia, X., and Ren, W. (2017). The temporal dynamics underlying the comprehension of scientific metaphors and poetic metaphors. *Brain Res.* 1655, 33–40. doi: 10.1016/j.brainres.2016.11.005

Tartter, V. C., Gomes, H., Dubrovsky, B., Molholm, S., and Stewart, R. V. (2002). Novel metaphors appear anomalous at least momentarily: evidence from N400. *Brain Lang.* 80, 488–509. doi: 10.1006/brln.2001.2610

Tesche, C. D., and Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *PNAS USA* 97, 919–924.

Tsur, R. (1992). *Toward a Theory of Cognitive Poetics*. Amsterdam, North Holland.

Weiss, P. H., Marshall, J. C., Wunderlich, G., Tellmann, L., Halligan, P. W., Freund, H.-J., et al. (2000). Neural consequences of acting in near versus far space: a physiological basis for clinical dissociations. *Brain* 123, 2531–2541. doi: 10.1093/brain/123.12.2531

Wlotko, E. W., and Federmeier, K. D. (2007). Finding the right word: hemispheric asymmetries in the use of sentence context information. *Neuropsychologia* 45, 3001–3014. doi: 10.1016/j.neuropsychologia.2007.05.013

Xie, M., Yang, Q., and Wang, Q. (2016). The P200 component in lexical processing. *Adv. Psychol.* 6, 114–120. doi: 10.12677/AP.2016.62015

Zhang, H., Jiang, L., Gu, J., and Yang, Y. (2013). Electrophysiological insights into the processing of figurative two-part allegorical sayings. *J. Neurolinguistics* 26, 421–439. doi: 10.1016/j.jneuroling.2013.01.004

Zhao, M., Meng, H., Xu, Z., Du, F., Liu, T., Li, Y., et al. (2011). The neuromechanism underlying verbal analogical reasoning of metaphorical relations: An event-related potentials study. *Brain Res.* 1425, 62–74. doi: 10.1016/j.brainres.2011.09.041

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 27 July 2022
doi: 10.3389/fpsyg.2022.916854

Check for updates

Metonymy Processing in Chinese: A Linguistic Context-Sensitive Eye-Tracking Preliminary Study

Xianglan Chen[1], Hulin Ren[2] and XiaoYing Yan[1]*

[1]*Center for the Cognitive Science of Language, Beijing Language and Culture University, Beijing, China,* [2]*School of Foreign Studies, University of Science and Technology, Beijing, China*

Current cognitively oriented research on metaphor proposes that understanding metaphorical expressions is a process of building embodied simulations, which are constrained by past and present bodily experiences. However, it has also been shown that metaphor processing is also constrained by the linguistic context but, to our knowledge, there is no comparable work in the domain of metonymy. As an initial attempt to fill this gap, the present study uses eye-tracking experimentation to explore this aspect of Chinese metonymy processing. It complements previous work on how the length of preceding linguistic context influences metonymic processing by focusing on: (1) the contextual information of both the preceding target words; (2) the immediate spillover after the target words; and (3) whether the logical relationship between the preceding contextual information and the target word is strong or weak (a 2 × 2 between-subject experiment with target words of literal/metonymy and logic of strong/weak). Results show that readers take longer to arrive at a literal interpretation than at a metonymic one when the preceding information is in a weak logic relationship with target words, although this disparity can disappear when the logic is strong. Another finding is that both the preceding and the spillover contextual information contribute to metonymy processing when the spillover information does more to the metonymy than it does to the literal meaning. This study further complements cognitive and pragmatic approaches to metonymy, which are centered on its conceptual nature and its role in interpretation, by drawing attention to how the components of sentences contribute to the metonymic processing of target words. Based on an experiment, a contextual model of Chinese metonymy processing is proposed.

Keywords: metonymy processing, eye tracking, spillover contextual information, embodied cognition, preceding contextual information

OPEN ACCESS

Edited by:
Laura M. Morett,
University of Alabama,
United States

Reviewed by:
Aseel Zibin,
The University of Jordan, Jordan
Francisco Ruiz De Mendoza,
University of La Rioja, Spain

****Correspondence:***
Hulin Ren
hulinr@aliyun.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *10 April 2022*
Accepted: *09 June 2022*
Published: *27 July 2022*

Citation:
Chen X, Ren H and Yan X (2022) Metonymy Processing in Chinese: A Linguistic Context-Sensitive Eye-Tracking Preliminary Study. Front. Psychol. 13:916854. doi: 10.3389/fpsyg.2022.916854

INTRODUCTION

In a recent cognitive-linguistic account, metonymy is described as a domain-internal conceptual mapping based on expansion or reduction cognitive operations which combine with a formal substitution operation (which is not exclusive of metonymy; *cf.* Ruiz de Mendoza, 2017; see also Ruiz de Mendoza Ibáñez, 2021, for an overview of positions). Ruiz de Mendoza Ibáñez and Pérez Hernández (2001) have further proposed the term *high-level metonymy* to refer to metonymies where both the source and target domains are abstract domains. An example is GENERIC FOR SPECIFIC (e.g., *What is that building*?, meaning the identity of the building, say, the Royal Palace). These and other cognitive-linguistic studies on metonymy (*cf.* Blanco Carrión et al., 2018) focus on its conceptual composition, its boundaries with other figures of speech (especially metaphor) and its interaction with other metonymies and with metaphor. Studies based on inferential pragmatics deal with the interpretive aspects of metonymy sometimes both in cognitive and communicative terms (e.g., Wilson and Carston, 2007; Jodłowiec and Piskorska, 2015). However, such concept-based linguistically-oriented proposals, no matter how developed, do not clearly take into account the influence of the preceding or following linguistic context on the target words. The situation is only slightly different in psycholinguistic research, which in addressing different aspects of metonymic processing, has incidentally touched on contextual factors. For example, anaphoric inference in figurative referential description was discussed by Gibbs (1990), who found that subjects were faster at reinstating the antecedents for literal referential descriptions than for metaphoric and metonymic descriptions and that people understood metaphoric referential descriptions more easily than they did metonymic descriptions. Duffy and Rayner (1990) proposed the influence of distance on figurative description, suggesting that the distance to the figurative description also affects interpretation. Frisson and Pickering (1999, 2001) used the literal context and the metonymic context to find that processing unfamiliar metonymies was more difficult than processing familiar ones. McElree et al. (2006) reported on an experiment that examined the processing of standard metonymies (e.g., *The gentleman read Dickens*) and logical metonymies (e.g., *The gentleman began Dickens*), which they contrasted with control expressions with a conventional interpretation (*The gentleman met Dickens*). Eye movement measures during the reading process indicated that logical metonymies were more costly to interpret. Experiments that have used preceding context that affects metonymic interpretation (e.g., Frisson and Pickering, 2007; Piñango et al., 2017; Chen and Li, 2019) have suggested that, only when the preceding context is adequate in terms of support for underspecification and the resolution of interpretive variability, is metonymic processing almost the same as the literal interpretation.

In the present study, we wondered if the context of preceding or following information could also affect metonymic processing. A few words of caution are necessary before we proceed to further contextualize and present our experimental work. First, such work is based on Chinese subjects. This means that our claims are valid only for the understanding of processing by native speakers of this language. To place this study fully within the context of others, based on English, it would be necessary to carry out a contrastive study, which is not our goal. However, the study we offer can be replicated for other languages in future work. Second, metonymy is a complex phenomenon. Scholars have distinguished lexical metonymy from predicative, predicational, and even illocutionary metonymy (Ruiz de Mendoza Ibáñez and Galera, 2014, p: 66; see also Panther, 2005, and Panther and Thornburg, 2018). Also, as noted above, even within the same level of metonymic activity (e.g., lexical) some metonymies involve more conceptual complexity than others (which is the case of metonymic chains; Barcelona, 2005; Brdar-Szabó and Brdar, 2011; *cf.* Ruiz de Mendoza Ibáñez and Galera, 2014, p: 117) and metaphor has long been found to interact with metonymy (Goossens, 1990). The present empirical study is an initial approximation to such a complex phenomenon. For this reason, we have chosen to discuss an indisputable and simple example of metonymy, 'scalpel for surgeon' (INSTRUMENT FOR USER). Again, future work should consider more complex analytical situations.

LITERATURE REVIEW

The Four Models of Figurative Language Processing

Since the beginning of figurative language research, many psycholinguistic studies (e.g., Inhoff et al., 1984; Gibbs, 1986, 1990; Keysar, 1989; Cronk et al., 1993; Onishi and Murphy, 1993; Schraw, 1995) have compared the temporal sequences involved in accessing the literal and figurative meanings of non-idiomatic (*He is an icebox*) and idiomatic metaphor (*He blew his stack*). Based on these comparisons, four models of figurative language processing have been proposed: the literal first model (Searle, 1979; Grice, 1989), the figurative first model (Gibbs, 1980; Estill and Kemper, 1982; Gibbs and Gonzales, 1985), the parallel model (Paivio, 1979; Glucksberg, 1991), and the underspecification model (Frisson and Pickering, 1999, 2001).

The literal first model is also known as the standard pragmatic view (Grice, 1975; Searle, 1979) or the indirect access model (Lowder and Gordon, 2013). In this model, access to the non-figurative meaning of a trope always precedes access to its figurative interpretation, which is adopted only when the literal meaning is rejected in relation to the context of the sentence. Although this model was supported by some early empirical studies (e.g., Janus and Bever, 1985), it was also challenged by the fact that, if sufficient context is provided, the processor accesses the figurative meaning first (Gibbs, 1980). Because of this, researchers put forward the figurative-first or direct access model (Gibbs and Gonzales, 1985), where subjects begins with a figurative interpretation when processing what appears to be figurative language but turn to the literal meaning when a figurative interpretation cannot be found or when the figurative interpretation is incongruous in the context of the sentence. Evidence for this model comes from psycholinguistic studies that have suggested that the figurative interpretation of an idiom is accessed before the non-figurative interpretation (Gibbs, 1986). The figurative-first model was later challenged by proponents of the parallel model (Glucksberg, 1991; Cacciari and Glucksberg, 1994), which asserts that, if both literal and figurative interpretations are appropriate in a context, neither is accessed prior to the other (Gerring and Healy, 1983; Myers et al., 1987). Similarly, Gerrig (1989) formulated the concurrent processing model, which maintains that the selection and creation of meaning in the processing of figurative language operate simultaneously. These three models of figurative language processing depend on a dichotomy between figurative and non-figurative meanings to compare the chronology of their access. Using a different focus, Frisson and Pickering (1999, 2001) proposed the underspecification model, which provides a more general explanation by dividing the processing of figurative description into two stages: (1) a schematic stage, in which an underspecified, schematic representation of the meaning of a figurative description is developed; and (2) a home-in stage, in which a specific, appropriate meaning of the figurative description is determined with the aid of contextual and lexical information. The present study explores metonymic processing in Chinese and establishes its own model, which is cognitive and contextual in nature, thus adding significantly to the pool of previous studies.

The embodiment hypothesis has a great influence on the understanding of figurative language, especially for metaphor interpretation. Challenging the view of the disembodied mind, the embodiment hypothesis holds that our conceptual system is based on and shaped by our bodily experience and the interactions with the environment (Lakoff and Johnson, 1999; Gibbs, 2006a). For example, while understanding metaphor, "people may create embodied simulations of speakers' messages that involve moment-by-moment 'what must it be like' processes which make use of ongoing tactile-kinesthetic experiences." (Gibbs, 2006b, p: 455). Nowadays, the notion of embodiment has been applied to explain various kinds of figurative language, such as metaphor, metonymy, and irony (Filik et al., 2015). The empirical findings in the present paper are consistent with the embodied view of figurative language (see "Discussion and Conclusion").

Eye-Tracking Studies and the Processing of Metonymy

Not many studies have addressed the processing of metonymy. Among them, there is work based on processing times (Gibbs, 1990), on offline comprehension tasks (Slabakova et al., 2013), self-pace reading and probe recognition (Zarcone et al., 2014), and reaction times for sensicality judgments (Bott et al., 2016). There are also a few studies based on eye-tracking technology (Frisson and Pickering, 1999; Lowder and Gordon, 2013; Chen and Li, 2019). There follows a brief overview of some of this latter work.

The subjects in Pickering and Traxler (1998) read a context sentence followed by a syntactically ambiguous target sentence that included a metonymy. The study showed that, although unfamiliar metonymic expressions were not interpreted as quickly as literal expressions, people could access them rapidly when they were given appropriate contextual information. Frisson and Pickering (1999) addressed this issue by comparing the processing of conventional and novel metonymies in their figurative and literal contexts using two eye-tracking experiments. They found that the processing of novel metonymies was more difficult than processing conventional metonymies in figurative contexts and that, compared with the literal context, the figurative context had a weak effect on people's ability to understand familiar metonymies. Their study highlighted the role of sentence context (or co-text) in understanding metonymies, especially unfamiliar ones. Frisson (2009) suggested that the underspecification model (Frisson and Pickering, 1999) prevailed in the processing of familiar metonymies in that readers begin their processing with a single, underspecified meaning instead of choosing between literal and figurative meanings. In a later study, Lowder and Gordon (2013) used eye-tracking experiments to determine the role of sentence structure in the processing of metonymy (Lowder and Gordon, 2013). The results indicated that the meanings of place-for-institution metonymies (e.g., *the White House* to refer to the American president and his advisers) were accessed more slowly than those metonymies which were the argument of a verb. The study also showed that, in the argument position, the metonymic expression required more time to process than ordinary nouns that refer to a person. These findings demonstrate that syntactic structure can modulate the processing of figurative language. Frisson and Pickering (2007) suggested that an adequate sentential context (or co-text) could facilitate the processing of a novel metonymy. Chen and Li (2019) also experimented with how the different lengths of preceding information affected the processing of metonymy in Chinese. The results showed that, with short sentence contexts, readers took longer to produce a metonymic interpretation than a literal one for unfamiliar metonymies. However, the processing disparity between metonymic comprehension and literal comprehension disappeared when more extensive and supporting information was available in the preceding sentence context.

In this context of empirical work, the present study takes up a still unexplored topic. It deals with how different levels of logic (strong/weak) work with the processing of literal/metonymic descriptions in Mandarin Chinese given contextual information from sentences of the same length. To be clear, a strong logic is one in which the information preceding the target words provides cause-effect evidence for its conclusion. For example, in the sentence *bingren taiduo, na ba* shoushudao *queshi shiyong tai pinfan le* (lit. Patients too many, that CL ***scalpel*** indeed use too frequently, where CL stands for classifier; 'Since there are too many patients, that scalpel is used too frequently indeed'), the surgeon needs to operate frequently **because of** the excess of patients. Our hypothesis is that it will be easier to understand the target words under strong logic conditions. By contrast, a weak logic is one that information before and after the target words shows no cause-effect relation. We hypothesize that it will be harder to understand the target words under such conditions. For example, in the sentence *Shiyanshi li na ba* shoushudao *queshi shiyong tai pinfan le* (laboratory in, that CL *scalpel* indeed use too frequently ASP, 'In that laboratory that scalpel is used too frequently'), the relationship between the text preceding *shoushudao* ('scalpel') and the text following it is not a cause-effect one.

To address this goal, we put forward the following research questions:

1. How does the preceding context facilitate the literal and metonymic description of the part-whole person metonymy type in Mandarin Chinese?
2. How does the level of logic (strong/leak) affect the relationship between the preceding context and the interpretation of the metonymic description of this metonymy type?
3. How does spillover text work in processing metonymy of this type?

MATERIALS AND METHODS

Participants

Forty students (15 males and 25 females, mean age 21.6, age range 18–27) majoring in linguistics at Peking University were recruited for the experiment. All were native speakers of Chinese and all had normal eyesight.

Materials

Eighty sets of sentences were constructed in Chinese. Each set had four types of conditions: (a) the literal-weak logic condition, in which the target word of the literal sense is used with a weak-logic preceding context; (b) the literal-strong condition, in which the target word is used in a literal sense with a strong-logic preceding context; (c) the metonymic-weak condition, in which the target word is used in a metonymic sense with aa weak-logic preceding context; and (d) the metonymic-strong condition, in which the target word is used in the metonymic sense with a strong-logic preceding context. The following set of sentences, with target words in bold, is an example of the material used in the experiment.

1 a. *Shiyanshi li, na ba shoushudao queshi shiyong tai pinfan le*.

laboratory in, that CL **scalpel** indeed use too frequently ASP.

'That scalpel in the laboratory was indeed used too frequently.'

b. *bingren taiduo，na ba shoushudao queshi shiyong tai pinfan le*.

patients too many, that CL **scalpel** indeed use too frequently ASP.

'Since there are so many patients, that scalpel was indeed used too frequently.'

c. *shiyanshi li, na wei shoushudao queshi zuotian mei lai zhudao.*

laboratory in, that CL **scalpel** indeed yesterday neg. Come operate.

'That scalpel did not come to perform the operation in the laboratory yesterday.'

d. *chuchai zai wai, na wei shoushudao queshi zuotian mei lai zhudao.*

on business in outside, that CL **scalpel** indeed yesterday neg. Come operate.

'Because (he) was on a business trip, that scalpel did not come to perform the operation yesterday.'

In Chinese, people use *shoushudao* (scalpel) either literally to refer to a tool, as in (a) and (b) or metonymically to refer to a doctor, as in (c) and (d). In 1(a) and 1(b), the word *ba* is used as a classifier to disambiguate the literal and metonymic meanings of *shoushudao*, as *ba* is an indicator of an object rather than a person. In 1(c) and (d), the word *wei* is used as a classifier to disambiguate the two possible senses, as *wei* is used only to modify people (rather than things). However, different levels of logic appear between the preceding context and the target words.

Pretest for the Experimental Material

To assess the appropriateness of our 80 sets of sentences before the pretest, 12 native Chinese speakers with PhD and MR degrees in psycholinguistics or linguistics were hired to screen out inappropriate sentences based primarily on intuition and to choose familiar metonymy and ensure literal or metonymic descriptions.

Non-experimental metonymy-identification procedures have been discussed in the recent literature on metonymy. Littlemore (2015, p: 127) puts forward a procedure adapted from the well-known metaphor-identification procedure in Steen (2007; see also Zibin, 2021, p: 33). In application of this procedure to one of our examples, *At school, this long-haired often travels with me*, the researcher first identifies metonymy-related words ('long-haired'), secondly propositions ('person who has long hair'), thirdly denoted entities (long hair and person), and finally, the researcher makes conceptual links between the denoted concepts and their associated conceptual domains: the fact that the hair is part of a person suggests a part-for-whole metonymic configuration where the long hair stands for the person characterized by having long hair. This procedure is systematic, but it relies on the analyst's knowledge, which can be considered subjective. In order to endow the identification of metonymies with intersubjectivity, which is akin to objectivity if supported by a reliable statistical procedure, we complemented the analyst's identification process with a questionnaire-based protocol described below.

Sixteen sets of sentences that were grammatically correct but pragmatically unreasonable were removed from the experimental material. Then the remaining 64 sets of sentences were randomized into four lists by means of a Latin square design for post-questionnaire using a five-point Likert scale. Two pretests were conducted to assess the sentences' acceptability in terms of effective selection of metonymies and the relevance of the preceding adverbial to the target word by means of a questionnaire that used a five-point Likert scale. The relevance test was performed to ensure that the distinction between the condition of weak logical contextual information and that of strong logical contextual information was clear. The results of the pretest on the acceptability of the experimental metonymic material showed a similar main effect of sense [F_2 (1,63) = 12.537, $p = 0.001$, $\eta^2 = 0.166$]. In the relevance test, the material that preceded the comma was perceived as being more logical in its relationship to the target word under the condition of strong logical information than it was under the condition of weak logical information ($M = 4.071$ versus 3.176, $p = 0.000$), as expected. In summary, the data confirmed that the selection of metonymies was effective and the manipulation of the relationship between the metonymies and the logical context was successful.

Finally, the 64 sets of sentences were divided into four lists using a Latin square design, with each set consisting of four conditions. To avoid the effect of stimulus predictability, we added

76 sentences to each list as fillers, and pseudo-randomization was applied to all lists.

Equipment and Procedure

We employed the eye tracker EyeLink 1,000, which uses infrared video-based tracking technology to sample the subjects' eyes at 1,000 Hz. The stimuli were programmed to be displayed on a computer screen about 50 centimeters away from the viewers. Participants took the test at a Peking University laboratory. Written instructions were distributed asking the participants not to move their heads or blink excessively during the experiment. Then the eye tracker was adjusted and calibrated, after which 10 practice items were presented, followed by a one-minute break before the real experiment began. Participants were asked to look at a black dot in the middle of the screen, followed by a crosshair at the exact point at which the first Chinese character in each sentence would appear. If the participant's eyes did not fix on the crosshair within a certain period of time, the black dot reappeared, as the test sentence would not appear until the participant's eyes were fixed on the crosshair. The participants were then asked to read the sentences and press a key to trigger a yes-no response to a question that checked their understanding of the test sentences. To keep a balance across the conditions, half of the comprehension questions required a "yes" answer, and the other half required a "no" answer. After answering the comprehension questions, participants moved on to the next trial. The entire experiment lasted approximately 50 min, divided into two sessions with a predetermined break in the middle. If the participant became tired or blinked too much during the experiment, the experiment was paused to allow the participant to rest.

FINDINGS

Eye-fixations that were shorter than 80 msec were removed from the data, as the participants could not have processed any information during such a short time (Henderson et al., 1989). Ninety-two trials (3.5%) in which the participants provided incorrect answers to the comprehension questions were removed from the study.

In the next phase of the test, four areas of interest (AOIs) were established in every sentence for observation: the adverbial in front of the comma, the classifier, the target word, and the spillover region. For the purpose of making clear the processing of each of the areas before and after the target word, six eye-movement measures were used in each of the four AOIs: *the duration of the first fixation* (the length of the first fixation on the current AOI); *the first-run dwell time* (the length of all fixations on the current AOI in the first run); *the second-run dwell time* (the length of all fixations on the current AOI in the second run); *regression in count* (the number of times the participant looked back at an AOI from a point later in the sentence); *duration of the regression path* (the time from when the first fixation on the AOI started to the end of the last fixation before leaving the AOI, which includes not only *first-run dwell time* of the region but also the rereading time of the previous regions); and *dwell time* (the sum of time spent on an AOI).

RESULTS

A general linear model repeated measures analysis was applied to analyze the eye-movement data. The results of ANOVAs are reported based on both the means for each subject across items (F_1) and the means for each item across subjects (F_2). Results from each critical region are discussed separately to represent the nature of the observed effects.

Target-Word Region

The target-word region refers to the literal or metonymic description in the sentence, which is the core part of the experiment. **Table 1** shows the mean and standard errors of the participants' six eye-movement measures.

As **Table 1** shows, two main effects on two eye-movement measures: regression in count and dwell time. Regression in count shows that a preceding context with weak logical information caused the target words to receive more regressions from a later area [$F_1(1,39) = 15.068$, $p = 0.000$, $\eta^2 = 0.279$; $F_2(1,63) = 11.648$, $p = 0.001$, $\eta^2 = 0.156$], while dwell time on target words was longer under the condition of weak logic than under that of strong logical information [$F_1(1,39) = 5.591$, $p = 0.023$, $\eta^2 = 0.125$; $F_2(1,63) = 4.809$, $p = 0.032$, $\eta^2 = 0.071$]. Both the item analysis and the subject analysis reveal a significant interaction in the duration of the first-run fixation [$F_1(1,39) = 9.116$, $p = 0.004$, $\eta^2 = 0.189$; $F_2(1,63) = 7.587$, $p = 0.008$, $\eta^2 = 0.107$]. Further analysis of this simple effect demonstrates that the duration of the first-run fixation was longer for target words with literal sense than it was for target words with metonymic sense when there was less information in the preceding context [$F_1(39) = 6.189$, $p = 0.017$; $F_2(63) = 4.207$, $p = 0.047$]. Meanwhile, when the target words required literal interpretation, they saw longer duration of fixation in the first run under the condition of weak logic than under the condition of strong logic [$F_1(39) = 6.256$, $p = 0.017$; $F_2(63) = 8.109$,

TABLE 1 | Mean reading times and standard errors for the target word.

Measures	Condition			
	Literal-less	Literal-more	Metonymic-less	Metonymic-more
Duration of first fixation	225.60(5.77)	214.87(4.68)	213.38(5.86)	220.40(5.60)
First-run dwell time	275.76(9.89)	262.79(9.33)	265.52(9.82)	273.85(10.37)
Regression in count	0.45(0.041)	0.38(0.041)	0.48(0.048)	0.41(0.045)
Duration of the regression path	320.91(14.17)	308.61(11.93)	304.81(16.31)	315.71(13.91)
Second-run dwell time	261.61(11.46)	256.18(11.58)	263.14(18.41)	251.74(10.13)
Dwell time	407.45(19.65)	369.54(19.77)	405.83(16.72)	397.61(18.17)

$p = 0.006$]. This result is supportive of the main effect mentioned above.

Adverbial Region

When we explore the adverbial region, which works as part of the preceding context, it is helpful to know the degree to which the preceding context functions in the processing of the key words. **Table 2** shows the mean of participants' eye-movement data regarding the adverbial region.

As **Table 2** shows, for the difference between literal and metonymic description, there is a significant main effect of sense in three measures [first-run dwell time: $F_1(1,39) = 6.116$, $p = 0.018$, $\eta^2 = 0.136$; $F_2(1,63) = 3.517$, $p = 0.065$, $\eta^2 = 0.053$; duration of the regression path: $F_1(1,39) = 7.067$, $p = 0.011$, $\eta^2 = 0.153$; $F_2(1,63) = 4.247$, $p = 0.043$, $\eta^2 = 0.063$; and dwell time: $F_1(1,39) = 10.738$, $p = 0.002$, $\eta^2 = 0.216$; $F_2(1,63) = 3.976$, $p = 0.050$, $\eta^2 = 0.059$], suggesting that participants spent more time reading the adverbial region when they were in the literal condition than when they were in the metonymy condition. This result is supported by another two measures—the duration of fixation duration and regression in count—which showed a marginally significant main effect of sense, although only in the analysis of the subject [duration of the first fixation: $F_1(1,39) = 3.277 p = 0.078$, $\eta^2 = 0.078$; $F_2(1,63) = 1.310$, $p = 0.257$, $\eta^2 = 0.020$; regression in count: $F_1(1,39) = 2.855$, $p = 0.099$, $\eta^2 = 0.068$; $F_2(1,63) = 1.217$, $p = 0.274$, $\eta^2 = 0.019$].

The preceding context had a significant main effect on the same three measures [first-run dwell time: $F_1(1,39) = 82.789$, $p = 0.000$, $\eta^2 = 0.680$; $F_2(1,63) = 41.764$, $p = 0.000$, $\eta^2 = 0.399$; duration of regression path: $F_1(1,39) = 82.871$, $p = 0.000$, $\eta^2 = 0.680$; $F_2(1,63) = 47.543$, $p = 0.000$, $\eta^2 = 0.430$; dwell time: $F_1(1,39) = 99.393$, $p = 0.000$, $\eta^2 = 0.718$; $F_2(1,63) = 34.573$, $p = 0.000$, $\eta^2 = 0.354$], as well as the regression count [$F_1(1,39) = 22.631$, $p = 0.000$, $\eta^2 = 0.367$; $F_2(1,63) = 15.810$, $p = 0.000$, $\eta^2 = 0.201$], revealing that participants spent more time reading the adverbial region when there was weak logic in the preceding information. Support for this result comes from two other measures: duration of the first fixation, which showed a marginally significant effect of preceding context only in the item analysis [$F_1(1,39) = 1.935$, $p = 0.172$, $\eta^2 = 0.047$; $F_2(1,63) = 2.964$, $p = 0.090$, $\eta^2 = 0.045$], and the second-run dwell time, which showed a significant effect of the preceding context only in the subject analysis [$F_1(1,39) = 5.587$, $p = 0.024$, $\eta^2 = 0.138$; $F_2(1,63) = 1.935$, $p = 0.169$, $\eta^2 = 0.032$].

TABLE 2 | Mean reading time for the adverbial region.

Measures	Condition			
	Literal-less	Literal-more	Metonymic-less	Metonymic-more
Duration of the first fixation	278.19(4.14)	288.22(5.44)	280.93(4.57)	279.12(4.94)
First-run dwell time	458.11(15.20)	558.90(23.45)	457.76(15.64)	520.97(19.06)
Regression in count	0.37(0.047)	0.50(0.052)	0.35(0.048)	0.44(0.053)
Duration of the regression path	459.65(15.06)	567.38(23.54)	459.66(16.33)	523.90(19.18)
Dwell time	588.92(29.37)	739.76(35.63)	587.30(28.29)	667.73(31.43)

The subject analysis and the item analysis also have a significant interaction effect on four of the six measures [duration of the first fixation: $F_1(1,39) = 4.457$, $p = 0.041$, $\eta^2 = 0.103$; $F_2(1,63) = 3.884$, $p = 0.053$, $\eta^2 = 0.058$; first-run dwell time: $F_1(1,39) = 12.951$, $p = 0.001$, $\eta^2 = 0.249$; $F_2(1,63) = 4.560$, $p = 0.037$, $\eta^2 = 0.067$; duration of the regression path: $F_1(1,39) = 16.420$, $p = 0.000$, $\eta^2 = 0.296$; $F_2(1,63) = 6.062$, $p = 0.017$, $\eta^2 = 0.088$; dwell time: $F_1(1,39) = 12.425$, $p = 0.001$, $\eta^2 = 0.242$; $F_2(1,63) = 4.253$, $p = 0.043$, $\eta^2 = 0.063$]. Additional analysis of these simple effects shows no difference in the time participants spent reading adverbials between the metonymic and literal sentences when weak logic was provided [i.e., condition (a) vs. condition (c); duration of first fixation: $F_1(39) = 0.596$, $p = 0.445$; $F_2(63) = 0.390$, $p = 0.534$; first-run dwell time: $F_1(39) = 0.016$, $p = 0.939$; $F_2(63) = 0.006$, $p = 0.900$; duration of the regression path: $F_1(39) = 0.035$, $p = 0.852$; $F_2(63) = 0.020$, $p = 0.888$; dwell time: $F_1(39) = 0.029$, $p = 0.865$; $F_2(63) = 0.010$, $p = 0.920$]. However, the difference between the literal and the metonymic sentences was significant when strong logical information was provided in the preceding context [i.e., condition (b) vs. condition (d); duration of the first fixation: $F_1(39) = 7.292$, $p = 0.010$; $F_2(63) = 4.574$, $p = 0.036$; first-run dwell time: $F_1(39) = 15.1$, $p = 0.017$; $F_2(63) = 5.991$, $p = 0.000$; duration of the regression path: $F_1(39) = 18.557$, $p = 0.000$; $F_2(63) = 7.326$, $p = 0.009$; dwell time: $F_1(39) = 19.848$, $p = 0.000$; $F_2(63) = 6.483$, $p = 0.013$]. In other words, when the context was strong logical, participants spent more time on the adverbial region when the target words required a literal interpretation than they did when the target words required a metonymic interpretation. In addition, when a literal interpretation of the target words was required, condition (a) and (b) showed a significant difference [duration of the first fixation: $F_1(39) = 5.545$, $p = 0.024$; $F_2(63) = 5.729$, $p = 0.020$; first-run dwell time: $F_1(39) = 67.005$, $p = 0.000$; $F_2(63) = 34.827$, $p = 0.000$; duration of the regression path: $F_1(39) = 72.718$, $p = 0.000$; $F_2(63) = 40.996$, $p = 0.020$; dwell time: $F_1(39) = 71.647$, $p = 0.000$; $F_2(63) = 28.694$, $p = 0.002$]: that is, participants spent more time on the adverbial region when strong logical information was provided in the preceding context when literal interpretation was required. The same result was found in three measures for the sentences that required metonymic interpretations [i.e., conditions (c) and (d); first-run dwell time: $F_1(39) = 53.026$, $p = 0.000$; $F_2(63) = 23.992$, $p = 0.020$; duration of the regression path: $F_1(39) = 48.116$, $p = 0.000$; $F_2(63) = 23.930$, $p = 0.020$; dwell time: $F_1(39) = 47.784$, $p = 0.000$; $F_2(63) = 10.879$, $p = 0.002$].

Classifier Region

We performed an analysis to determine the contribution of the classifier region to metonymic processing key words. **Table 3** shows the average time the participants spent reading the classifier region (including a demonstrative and a classifier) under each condition for the eye movements measures.

As **Table 3** shows, the preceding context has a significant main effect on two eye-movement measures: regression in count [$F_1(1,39) = 3.778$, $p = 0.059$, $\eta^2 = 0.090$; $F_2(1,63) = 5.624$, $p = 0.021$, $\eta^2 = 0.083$] and duration of the regression path [$F_1(1,39) = 10.137$,

TABLE 3 | Reading times of the classifier region.

Measures	Condition			
	Literal-less	Literal-more	Metonymic-less	Metonymic-more
Duration of the first fixation	222.26(8.84)	228.180(8.80)	223.86(7.19)	230.79(6.95)
First-run dwell time	244.380(13.29)	249.39(13.27)	241.94(10.83)	249.99(9.97)
Regression in count	0.49(0.048)	0.4(0.048)	0.48(0.049)	0.44 (0.066)
Duration of the regression path	268.33(14.56)	287.48(14.71)	260.31(10.49)	281.58(11.87)
Dwell time	186.73(20.99)	207.40(22.00)	198.97(21.24)	193.17(20.17)

$p = 0.003$, $\eta^2 = 0.211$; $F_2(1,63) = 3.629$, $p = 0.061$, $\eta^2 = 0.055$], so the classifier region attracted more regression time from the later areas of target words when strong, instead of weak, logic was provided. The subject analysis of two other measures supported the result too. The duration of the first fixation and the first-run dwell time also presented a significant or marginally significant main effect of the preceding context [duration of the first fixation: $F_1(1,39) = 5.538$, $p = 0.024$, $\eta^2 = 0.127$; $F_2(1,63) = 1.306$, $p = 0.258$, $\eta^2 = 0.021$; first-run dwell time: $F_1(1,39) = 2.899$, $p = 0.097$, $\eta^2 = 0.071$; $F_2(1,63) = 1.170$, $p = 0.283$, $\eta^2 = 0.019$], demonstrating that the classifier region received more attention initially, when strong logical information was in the preceding context.

Spillover Region

The spillover region is explored to determine if this part of the sentence context contributes to the keywords of the metonymy. This examination is the first for this part of the context, but the Chinese language is not as linear as English, so further exploration is needed. **Table 4** shows the means of the participants' six eye-movement measures in the spillover region (with at least two Chinese characters immediately following the target word in each sentence) under the four conditions.

Table 4 shows that the spillover region has a significant main effect on the dwell time [$F_1(1,39) = 7.462$, $p = 0.009$, $\eta^2 = 0.161$; $F_2(1,63) = 5.084$, $p = 0.028$, $\eta^2 = 0.075$], indicating that participants took less time to read the spillover region if the target words required a literal interpretation than when these words required a metonymic interpretation.

Meanwhile, two measures indicating early processing, duration of the first fixation and first-run dwell time, reveal a significant interaction effect [duration of the first fixation: $F_1(1,39) = 4.009$, $p = 0.052$, $\eta^2 = 0.093$; $F_2(1,63) = 5.435$, $p = 0.023$, $\eta^2 = 0.081$; first-run dwell time: $F_1(1,39) = 4.812$, $p = 0.034$, $\eta^2 = 0.110$; $F_2(1,63) = 5.646$, $p = 0.021$, $\eta^2 = 0.083$]. Further study of this simple effect reveals that, when the preceding context contains weak logical information, conditions (a) and (c) are significantly different [duration of the first fixation: $F_1(39) = 2.865$, $p = 0.098$; $F_2(63) = 4.996$, $p = 0.029$; first-run dwell time: $F_1(39) = 5.165$, $p = 0.029$; $F_2(63) = 5.362$, $p = 0.024$], which shows that the spillover region receives less reading time if a literal interpretation of the target words, rather than the metonymic interpretation, is required. However, if more information is included in the preceding context, the difference between the literal condition (b) and the metonymic condition (d) becomes less significant [duration of the first fixation: $F_1(39) = 1.800$, $p = 0.187$; $F_2(63) = 1.566$, $p = 0.215$; first-run dwell time: $F_1(39) = 1.340$, $p = 0.254$; $F_2(63) = 1.333$, $p = 0.253$]. In addition, when the target words require a literal interpretation, the difference in the time the participants spent reading the spillover region between the logical context conditions (a) and (b) was insignificant [duration of the first fixation: $F_1(39) = 1.874$, $p = 0.179$; $F_2(63) = 3.139$, $p = 0.081$; first-run dwell time: $F_1(39) = 3.092$, $p = 0.087$; $F_2(63) = 5.362$, $p = 0.024$]. However, when a metonymic interpretation was required, the difference between conditions (c) and (d), where the spillover region received a longer first fixation when weak logical information was provided in the preceding context, was marginally significant [$F_1(39) = 3.310$, $p = 0.077$; $F_2(63) = 2.606$, $p = 0.112$].

TABLE 4 | Mean reading time of the spillover region.

Measures	Condition			
	Literal-less	Literal-more	Metonymic-less	Metonymic-more
Duration of the first fixation	219.48(4.78)	227.58(6.36)	231.20(6.30)	221.53(5.25)
First-run dwell time	242.78(7.66)	257.13(10.51)	262.25(9.97)	249.67(8.56)
Regression in count	0.30(0.031)	0.23(0.03)	0.29(0.035)	0.30(0.035)
Duration of the regression path	311.68(14.10)	317.20(16.01)	329.42(14.88)	330.098(16.40)
Second-run dwell time	237.26(8.71)	238.89(11.35)	241.13(11.95)	239.17(11.12)
Dwell time	296.57(16.86)	297.40(15.44)	327.59(17.45)	316.86(16.03)

DISCUSSION AND CONCLUSION

This study differs from other studies on metonymy in that it considers only the logical relationship between the preceding sentence context and the target words. Ortony et al. (1978) proposed that preceding text inspired readers to form a framework for understanding the figurative language, but the present study goes further to show how regions of both the preceding sentence context and the following sentence context (the spillover region) contribute to literal and metonymic processing.

The results show that target words are processed faster when the preceding context contains strong, rather than weak, logic in relation to the key words, regardless of whether the target words require metonymic or literal interpretation. This outcome indicates that the meaning of the target word comes not only from the target word itself, but also from the reasoning from regression pass duration with regression counts. To be more exact, the time spent interpreting the two literal conditions [conditions (a) and (b)] is longer than that for the two metonymic conditions [conditions (c) and (d)]. The time spent in condition

(a) was much longer than that in condition (b), suggesting that strong logical information in the preceding context (b) facilitates the processing of literal senses. Literally used target words in Mandarin Chinese were more sensitive to the preceding logical information, but the disparity in reading time between literal and metonymic meanings could be reduced if strong logical information were available in the preceding context. Metonymic target words depend more on the classifier than the logical information of the previous context, as illustrated in **Figure 1**.

As for the measure of the duration of the regression path in the adverbial region, the logic in the preceding context facilitates the processing of both the literal and metonymic target words, although the logic works more effectively in the literal than in the metonymic meaning. The dwell time offers the same support, although the classifier region and the spillover with the regression time contribute more to metonymy processing.

In this experiment, Frisson and Pickering's (2001) underspecification model is most likely to provide an explanation for the processing of metonymy, where meaning is underspecified in that the initial interpretations are compatible with many senses instead of only a certain sense. Based on contextual information, readers can then refine the meaning to fit a particular sense. If the preceding context is helpful in determining the appropriate sense, readers may home in on the specific sense more rapidly. The experiment conducted here shows that metonymic processing requires more contribution from the classifier that precedes it and the spillover that follows it than does literal processing.

This study also suggests that people's interpretation of metonymy partly involved creating an embodied simulation along with contextual information. Even though relationships are not physical entities that literally travel along physical paths, people nonetheless conceive of relationships in metaphorical ways, especially when prompted to do so by contextual information. This metonymic conceptualization is not purely abstract, but embodied in the sense that participants imagine themselves in the situation and experience. Of course, the study above used a specific sentence, and it is also unclear whether these findings generalize to other kinds of figurative language, including different types of metonymy (Gibbs, 2006b).

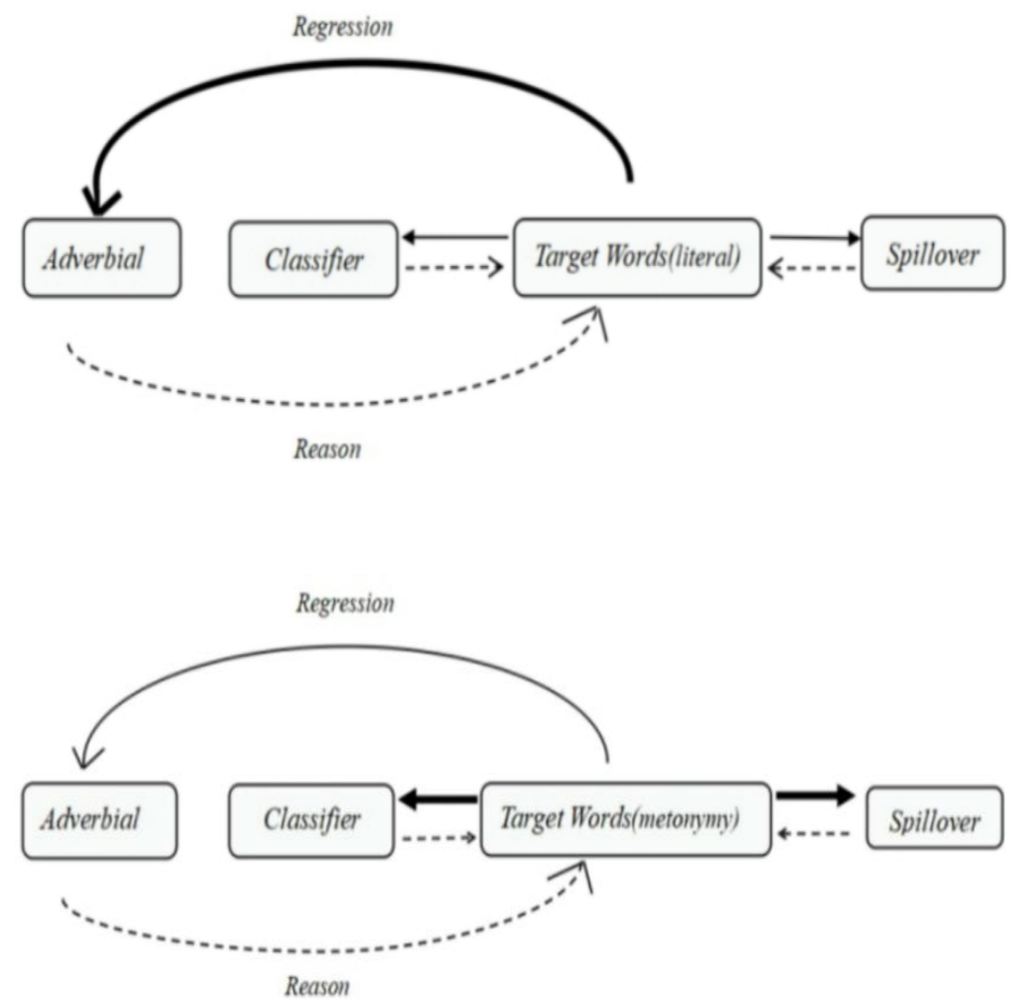

FIGURE 1 | Cognitive-psychological model for Chinese metonymy as subject.

The findings of the study support the notion that familiar metonymy may result in faster processing. The results of the eye-tracking experiment also provide strong evidence for the contribution of the different regions to interpretation. Future research is needed to determine how the classifier (CL) contributes to processing the target words.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XC provides the idea and design and writes the draft paper. HR is in charge of the experiment. XY gets the data and does the analysis. All authors contributed to the article and approved the submitted version.

FUNDING

We are grateful for the financial support provided by National Philosophy and Social Science Fund Program, number: 19BYY016 and the Project Funded by Ministry of Education, number: 19JHQ035 and University Project KCSZ202123.

REFERENCES

Barcelona, A. (2005). "The multilevel operation of metonymy in grammar and discourse, with particular attention to metonymic chains," in *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction*. eds. F. J. Ruiz de Mendoza and S. Peña (Berlin & New York: Mouton de Gruyter), 313–352.

Blanco Carrión, O., Barcelona, A., and Pannain, R. (Eds.) (2018). *Conceptual Metonymy. Methodological, Theoretical, and Descriptive Issues*. Amsterdam & Philadelphia: John Benjamins, 27–54.

Bott, L., Rees, A., and Frisson, S. (2016). The time course of familiar metonymy. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 1160–1170. doi: 10.1037/xlm0000218

Brdar-Szabó, R., and Brdar, M. (2011). "What do metonymic chains reveal about the nature of metonymy?" in *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*. eds. R. Benczes, A. Barcelona and F. J. Ruiz de Mendoza (Amsterdam & Philadelphia: John Benjamins), 217–248.

Cacciari, C., and Glucksberg, S. (1994). "Understanding figurative language," in *Handbook of Psycholinguistics*. ed. M. A. Gernsbacher (San Diego, CA: Academic Press), 447–477.

Chen, X., and Li, F. (2019). The length of preceding context influences metonymy processing: Evidences from eye tracking experiment. *Rev. Cogn. Linguist.* 17, 243–256.

Cronk, B., Lima, S., and Schweigert, W. (1993). Idioms in sentences: effects of frequency, literalness, and familiarity. *J. Psycholinguist. Res.* 22, 59–82. doi: 10.1007/BF01068157

Duffy, S., and Rayner, K. (1990). Eye movements and anaphor resolution: effects of antecedent typicality and distance. *Lang. Speech* 33, 103–119. doi: 10.1177/002383099003300201

Estill, R., and Kemper, S. (1982). Interpreting idioms. *J. Psycholinguist. Res.* 11, 559–568. doi: 10.1007/BF01067612

Filik, R., Hunter, C. M., and Leuthold, H. (2015). When language gets emotional: irony and the embodiment of affect in discourse. *Acta Psychol.* 156, 114–125. doi: 10.1016/j.actpsy.2014.08.007

Frisson, S. (2009). Semantic underspecification in language processing. *Lang. Ling. Compass* 3, 111–127. doi: 10.1111/j.1749-818X.2008.00104.x

Frisson, S., and Pickering, M. (1999). The processing of metonymy: evidence from eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1366–1383. PMID: 10605827

Frisson, S., and Pickering, M. (2001). Obtaining a figurative interpretation of a word: support for underspecification. *Metaphor. Symb.* 16, 149–171. doi: 10.1080/10926488.2001.9678893

Frisson, S., and Pickering, M. (2007). The processing of familiar and novel senses of a word: why reading Dickens is easy but reading Needham can be hard. *Lang. Cogn. Process.* 22, 595–613. doi: 10.1080/01690960601017013

Gerrig, R. (1989). The time course of sense creation. *Mem. Cogn.* 17, 194–207. doi: 10.3758/BF03197069

Gerring, R., and Healy, A. (1983). Dual processes in metaphor understanding: comprehension and appreciation. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 667–675. doi: 10.1037/0278-7393.9.4.667

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Mem. Cogn.* 8, 149–156. doi: 10.3758/BF03213418

Gibbs, R. W. (1986). Skating on thin ice: literal meaning and understanding idioms in conversation. *Discourse Process.* 9, 17–30. doi: 10.1080/01638538609544629

Gibbs, R. W. (1990). Comprehending figurative referential descriptions. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 56–66. doi: 10.1037/0278-7393.16.1.56

Gibbs, R. W. (2006a). *Embodiment in Cognitive Science*. New York: Cambridge University Press.

Gibbs, R. W. (2006b). Metaphor interpretation as embodied simulation. *Mind Lang.* 21, 434–458. doi: 10.1111/j.1468-0017.2006.00285.x

Gibbs, R. W., and Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition* 20, 243–259. doi: 10.1016/0010-0277(85)90010-1

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychol. Sci.* 2, 146–152. doi: 10.1111/j.1467-9280.1991.tb00122.x

Goossens, L. (1990). Metaphtonymy: The interaction of metaphor and metonymy in expressions for linguistic action. *Cogn. Linguist.* 1, 323–342. doi: 10.1515/cogl.1990.1.3.323

Grice, H. P. (1975). "Logic and conversation," in *Syntax and Semantics*. *vol. 3.* eds. P. Cole and J. Morgan, (New York: Academic Press), 41–58.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Henderson, J. M., Pollatsek, A., and Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Percept. Psychophys.* 45, 196–208. doi: 10.3758/BF03210697

Inhoff, A. W., Lima, S. D., and Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Mem. Cogn.* 12, 558–567. doi: 10.3758/BF03213344

Janus, R. A., and Bever, T. G. (1985). Processing of metaphoric language: An investigation of the three-stage model of metaphor comprehension. *J. Psycholinguist. Res.* 14, 473–487. doi: 10.1007/BF01666722

Jodłowiec, M., and Piskorska, A. (2015). Metonymy revisited: towards a new relevance-theoretic account. *Intercult. Pragmat.* 12, 161–187. doi: 10.1515/ip-2015-0009

Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *J. Mem. Lang.* 28, 375–385. doi: 10.1016/0749-596X(89)90017-X

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.

Littlemore, J. (2015). *Metonymy. Hidden Shortcuts in Language, Thought, and Communication*. Cambridge: Cambridge University Press.

Lowder, M. W., and Gordon, P. C. (2013). It's hard to offend the college: effects of sentence structure on figurative-language processing. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 993–1011. doi: 10.1037/a0031671

McElree, B., Murphy, G. L., and Ochoa, T. (2006). Time course of retrieving conceptual information: A speed-accuracy trade-off study. *Psychon. Bull. Rev.* 13, 848–853. doi: 10.3758/BF03194008

Myers, J. L., Shinjo, M., and Duffy, S. A. (1987). Degree of causal relatedness and memory. *J. Mem. Lang.* 26, 453–465. doi: 10.1016/0749-596X(87)90101-X

Onishi, K. H., and Murphy, G. L. (1993). Metaphoric reference: when metaphors are not understood as easily as literal expressions. *Mem. Cogn.* 21, 763–772. doi: 10.3758/BF03202744

Ortony, A., Schallert, D. L., Reynolds, R. E., and Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *J. Verbal Learn. Verbal Behav.* 17, 465–477. doi: 10.1016/S0022-5371(78)90283-9

Paivio, A. (1979). "Psychological processes in the comprehension of metaphor," in *Metaphor and Thought*. ed. A. Ortony (Cambridge, England: Cambridge University Press), 150–171.

Panther, K.-U. (2005). "The role of conceptual metonymy in meaning construction," in *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction*. eds. F. J. Ruiz de Mendoza and S. Peña (Berlin & New York: Mouton de Gruyter), 353–386.

Panther, K.-U., and Thornburg, L. (2018). "What kind of reasoning mode is metonymy?" in *Conceptual Metonymy. Methodological, Theoretical, and Descriptive Issues*. eds. O. Blanco Carrión, A. Barcelona and R. Pannain (Amsterdam & Philadelphia: John Benjamins), 121–160.

Pickering, M. J., and Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 940–961. doi: 10.1037/0278-7393.24.4.940

Piñango, M. M., Zhang, M., Foster-Hanson, E., Negishi, M., Lacadie, C., and Constable, R. T. (2017). Metonymy as referential dependency: psycholinguistic and neurolinguistic arguments for a unified linguistic treatment. *Cogn. Sci.* 41, 351–378. doi: 10.1111/cogs.12341

Ruiz de Mendoza, F. J. (2017). "Metaphor and other cognitive operations in interaction: From basicity to complexity," in *Metaphor: Embodied Cognition, and Discourse*. ed. B. Hampe (Cambridge: Cambridge University Press), 138–159.

Ruiz de Mendoza Ibáñez, F. (2021). "Conceptual metonymy theory revisited. Some definitional and taxonomic issues," in *The Routledge Handbook of Cognitive Linguistics*. eds. X. Wen and J. Taylor (London and New York: Routledge), 204–227.

Ruiz de Mendoza Ibáñez, F., and Galera, A. (2014). *Cognitive Modeling. A Linguistic Perspective*. Amsterdam & Philadelphia: John Benjamins.

Ruiz de Mendoza Ibáñez, F., and Pérez Hernández, L. (2001). Metonymy and the grammar: motivation, constraints and interaction. *Lang. Commun.* 21, 321–357. doi: 10.1016/S0271-5309(01)00008-8

Schraw, G. (1995). Components of metaphoric processing. *J. Psycholinguist. Res.* 24, 23–38. doi: 10.1007/BF02146098

Searle, J. R. (1979). *Expression and Meaning: Studies in the Theory of Speech Acts*. United Kingdom: Cambridge University Press.

Slabakova, R., Cabrelli Amaro, J., and Kang, S. K. (2013). Regular and novel metonymy in native Korean, Spanish, and English: experimental evidence for various acceptability. *Metaphor. Symb.* 28, 275–293. doi: 10.1080/10926488.2013.826556

Steen, G. J. (2007). Finding metaphor in discourse: pragglejaz and beyond. *Cult. Lang. Represent.* 5, 9–25.

Wilson, D., and Carston, R. (2007). "A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts," in *Pragmatics*. ed. N. Burton-Roberts (Basingstoke: Palgrave Macmillan), 230–259.

Zarcone, A., Padó, S., and Lenci, A. (2014). Logical metonymy resolution in a words-as-cues framework: evidence from self-paced reading and probe recognition. *Cogn. Sci.* 38, 973–996. doi: 10.1111/cogs.12108

Zibin, A. (2021). Blood metaphors and metonymies in Jordanian Arabic and English. *Rev. Cogn. Linguist.* 19, 26–50. doi: 10.1075/rcl.00075.zib

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fpsyg.2022.912176

Check for updates

OPEN ACCESS

EDITED BY
Huili Wang,
Dalian University of Technology, China

REVIEWED BY
Jingjing Guo,
Shaanxi Normal University, China
Wenbo Luo,
Liaoning Normal University, China

*CORRESPONDENCE
Zhao Yao
yaozhao@mail.xjtu.edu.cn
Fei Wang
fwang@xjtu.edu.cn

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 04 April 2022
ACCEPTED 18 July 2022
PUBLISHED 02 September 2022

CITATION
Yao Z, Chai Y, Yang P, Zhao R and Wang F (2022) Effects of social experience on abstract concepts in semantic priming. *Front. Psychol.* 13:912176.
doi: 10.3389/fpsyg.2022.912176

Effects of social experience on abstract concepts in semantic priming

Zhao Yao[1]*, Yu Chai[1], Peiying Yang[2], Rong Zhao[1] and Fei Wang[1]*

[1]School of Foreign Studies, Xi'an Jiaotong University, Xi'an, China, [2]School of Humanities, Xidian University, Xi'an, China

Humans can understand thousands of abstract words, even when they do not have clearly perceivable referents. Recent views highlight an important role of social experience in grounding of abstract concepts and sub-kinds of abstract concepts, but empirical work in this area is still in its early stages. In the present study, a picture-word semantic priming paradigm was employed to investigate the contribution effect of social experience that is provided by real-life pictures to social abstract (SA, e.g., *friendship, betrayal*) concepts and emotional abstract (EA, e.g., *happiness, anger*) concepts. Using a lexical decision task, we examined responses to picture-SA word pairs (Experiment 1) and picture-EA word pairs (Experiment 2) in social/emotional semantically related and unrelated conditions. All pairs shared either positive or negative valence. The results showed quicker responses to positive SA and EA words that were preceded by related vs. unrelated prime pictures. Specifically, positive SA words were facilitated by the corresponding social scene pictures, whereas positive EA words were facilitated by pictures depict the corresponding facial expressions and gestures. However, such facilitatory effect was not observed in negative picture-SA/EA word conditions. This pattern of results suggests that a facilitatory effect of social experience on abstract concepts varies with different sub-kinds of abstract concepts, that seems to be limited to positive SA concepts. Overall, our findings confirm the crucial role of social experience for abstract concepts and further suggest that not all abstract concepts can benefit from social experience, at least in the semantic priming.

KEYWORDS

abstract concepts, social experience, semantic priming, lexical processing, types of concepts

Introduction

Abstract concepts (e.g., *freedom*) are cognitively more complex compared with concrete concepts (e.g., *cat*) because they do not possess a bounded, identifiable object as referent, thus, their perceived content is more variable both within and across individuals. The way in which abstract concepts are acquired and represented has become a topic of intense debate in recent years, especially after the emergence of the embodied approaches to cognition (Yee, 2019). Within this framework, an integrate embodied view, multiple representation theory was proposed, suggesting that abstract concepts could be grounded in sensorimotor systems (e.g., visual and motor information) like concrete concepts, but they would activate to a larger extent linguistic, emotional, and social experience (Borghi et al., 2017, 2019; Pecher, 2018). Previous studies have provided plenty of behavioral and neurophysiological evidence on the role of linguistic experience (Louwerse and Jeuniaux, 2010) and affective experience (Kousta et al., 2011; Vigliocco et al., 2014; Yao et al., 2019) in grounding of abstract concepts. However, the role of social experience in abstract concepts has not received adequate attention (Davis et al., 2020; Fini et al., 2021).

Until recently, the Words As social Tools (WAT) view explicitly emphasizes the importance of social experience for abstract concepts, proposing that social experience is a constitutive part of abstract concepts (Borghi et al., 2019). Previous studies that used a feature listing or property generation task reported that participants were more likely to list communicative acts, social actions, or feelings as properties for abstract concepts than for concrete words (Wiemer-Hastings and Xu, 2005; Recchia and Jones, 2012). Neurophysiological studies also found that abstract concepts could activate brain regions underlying social cognition (e.g., medial prefrontal cortex, superior temporal sulcus; Wilson-Mendenhall et al., 2011; Roversi et al., 2013; Wang and Bi, 2019).

Evidence from different cognition tasks has shown that abstract concepts processing can involve more sociality features (Pecher, 2018). For example, Fini et al. (2021) employed human-avatar motor interaction and concept guessing tasks and found that participants needed more hints from partners in order to guess abstract concepts. Zdrazilova et al. (2018) used a taboo task in which participants were asked to communicate words' meanings to a partner without using the words *pre se*. By analyzing verbal and gestural data, they found that participants' speech referenced more people and introspections during communicating the meanings of abstract words.

Moreover, an important advance on the grounding of abstract concepts is the recognition that they are not a unitary whole, but composed of different sub-categories of abstract concepts exist (Ghio et al., 2016; Mkrtychian et al., 2019; Villani et al., 2019). Several studies that used a feature production task (Harpaintner et al., 2018) or performed meta-analyses (Desai et al., 2018) suggested that abstract concepts can be quite different from one another in terms of the features they activate, such as numerical, emotional, moral, and theory of mind. Neurophysiological evidence also showed that specific brain responses were separately engaged for social, emotional, and numerical concepts processing (Moseley et al., 2012; Mellem et al., 2016; Dreyer and Pulvermüller, 2018; Bechtold et al., 2019). In short, there is growing interest for considering the category-specific approach that has been applied to research on concrete concepts (e.g., *animal, fruit, furniture*) to examine of fine-grained abstract categories.

Based on recent work on the role of social experience in the grounding of abstract concepts and their categories, it is uncontroversial that social experience is crucial for understanding and processing of abstract concepts, but it is not clear whether it plays the same role in processing different categories of abstract concepts. Some studies have indicated that abstract concepts could be grounded in different aspects of embodied dimensions (e.g., sensorimotor, social, and affective; Ghio et al., 2016; Borghi et al., 2019; Wang and Bi, 2019; Kiefer and Harpaintner, 2020). For instance, an exploratory analysis by Connell et al. (2018) reported that emotional concepts appear to rely more on inner affective experience than non-emotional concepts. A rating study that performed a cluster analysis revealed four categories of abstract concepts, including philosophical emotional, social, and physical concepts. Among them, philosophical concepts were more abstract than the others; physical concepts were more concrete and more linked to interactions with external environment; by comparison, emotional concepts were more characterized by inner grounding, and social concepts relied both on inner and external grounding (Villani et al., 2019). These studies imply that although emotional and social concepts are associated only with inner affective information, but social concepts are grounded by both inner affective and social information.

Given all that, the aim of the present study is to examine whether social experience plays a specific role in the grounding of Social Abstract (SA) concepts, relative to Emotional Abstract (EA) concepts. The core question is whether social experience could be a main embodied dimension for characterizing SA concepts and effectively distinguishing them from other categories of abstract concepts. In the current study, SA concepts refer to general social knowledge that emerges from interpersonal interactions (e.g., *betrayal, duty, loyalty*; Bolognesi and Steen, 2018; Wang and Bi, 2019) and generally have an affective connotation (Giffard and Pratique, 2015). EA concepts refer to basic emotional experience that can directly label individual's internal affective states (e.g., *happiness, anger, sad*; Kazanas and Altarriba, J., 2015). For example, the SA word "*friendship*" contains a social knowledge of *"a closing and lasting relation between you and a person you like"* and a positively affective experience such as "*happiness*." In contrast, the EA word "*sad*" only conveys a negative feeling such as "*unhappiness.*"

Although SA and EA concepts lack clear and perceivable referents, they still evoke specific social scenes, episodes, or affective states and in turn are represented in terms of message from pictures depicting people, places, objects, and an individual's internal state (Barsalou et al., 2008; Wilson-Mendenhall et al., 2011). For instance, the social scene of *two boys with a smile have a snow fight in the park* might help us to understand the meaning of "*childhood.*" By contrast, the meaning of the EA word "*cheerfulness*" could be obtained from *boys' smiling face and body posture*. Concrete elements, such as *boys, snow fight, smiling faces, and park*, construct a particular social scene that could be experienced first-hand or could be heard or seen and thus effectively represent meanings of abstract concepts. Importantly, these concrete elements of pictures are usually perceived as a whole (Wilson-Mendenhall et al., 2011; McRae et al., 2018; Pecher, 2018). From this perspective, it is believable that pictures of social scenes, facial displays or body gestures automatically trigger relevant concept representation.

In the present study, we employed a "picture-word" semantic priming paradigm in a lexical decision task to compare the contributions of social experience to SA concepts and EA concepts. Semantic priming is a typical paradigm used to examine mental representations of word meanings and their relationships (Meyer and Schvaneveldt, 1971). Semantic priming effect refers to the faster and more accurate response to a target (e.g., *love*) when it is preceded by a semantically related prime (e.g., *marry*), relative to an unrelated prime (e.g., *news*). This effect is usually explained by an automatic spread of activation through semantic memory (Neely, 1991). Because pictures, similar to words, have a direct and functional connection to semantic system, with a similar activation like words in semantic memory (Glaser, 1992; Herring et al., 2013), and thus this paradigm could provide a means of measuring semantic relationship of picture-word pairs.

We conducted two experiments to examine responses to picture-SA word pairs (Experiment 1) and picture-EA word pairs (Experiment 2) in semantically related and unrelated conditions, in which all pairs shared either positive or negative valence. We considered stimuli valence is due to the fact that, in the literature on semantic priming, the difference in priming effects has been found between positive and negative primes (Rossell and Nobre, 2004). Therefore, four types of prime pictures were used, two of them were Social Scene (SS) pictures with either positive or negative valence, describing people's social interactions in a real-life situation, and the other two were positive or negative Emotional Expression (EE) pictures, describing a person's facial expression and gesture but without social interaction and situation. Accordingly, SA and EA target words were, respectively, presented under four experimental conditions in Experiment 1 and 2, including social-semantic primes (i.e., SS pictures) and/or non-social semantic primes (i.e., EE pictures), and all prime-target pairs shared either positive or negative valence.

We hypothesized that in Experiment 1, if social experience is a more significant factor for SA concepts than for EA concepts, then SA target words would be more readily facilitated by social-semantic related SS pictures, compared to social-semantic unrelated EE pictures. Conversely, if EA concepts are more detached form social experience or more dependent on inner affective states than SA concepts, then in Experiment 2, EA target words would be more readily facilitated by emotional-semantic related EE pictures, relative to social-semantic related SS pictures. That is, in the lexical priming-decision task, the effect of social experience provided by a real-life situational picture on the processing of SA and EA words might be an opposite. Meanwhile, the effect of social experience on SA and EA words might be modulated by their valence.

Norming study

We began with a norming study to generate stimuli for the subsequent experiments. A set of positive and negative picture-word pairs were created, consisting of social-semantic related pairs and emotional-semantic related pairs. This stage included four steps: (1) select pictures in term of the definitions of SS and EE pictures; (2) name each picture by reference to the method used in McRae et al. (2018); (3) rate semantic relationship between picture and its name; and (4) match affective and/or lexical variables of pictures and words.

Selected SS and EE pictures according to their definitions

We referred to the social–emotional sentences in Mellem et al. (2016) and initially collected 94 SS pictures and 96 EE pictures (half positive and half negative) from International Affective Picture System (IAPS; Lang et al., 2005), the Chinese Affective Picture System (CAPS, Bai et al., 2005) and public photos from the Internet. The selection of pictures was guided by the standard that a picture has a positive or negative affective meaning, and without conspicuous letters. We believed that these pictures convey the meaning of the corresponding words and can elicit consistent responses among participants. Two examples are shown in Figure 1.

A picture-naming task

Participants

Forty undergraduate students (26 males; mean age = 22.4, *SD* = 2.3) were recruited from Xidian University and were asked to provide 2–5 words related to each picture by an online questionnaire survey (SO JUMP, https://www.wjx.cn/). Each participant received 50 RMB for their participation.

Procedure and result

The procedure borrowed the method that was used in McRae et al. (2018)'s norming study. Participants were instructed to avoid naming people and objects, but rather to provide words summarizing the whole situation, or the affective states, or thoughts of the people in the picture. Appendix A presents the instructions. The task was self-paced, with all participants completing the task in 3 days.

According to a weighted score of each picture, we got the most appropriate name for each picture. The weighted score was calculated as follows: Rank Sum Score = 5a + 4b + 3c + 2d + e, where a, b, c, d, and e refer to the number of participants who provided that response in ranks 1, 2, 3, 4, and 5, respectively (McRae et al., 2018). The maximum possible rank sum score was 200 (40 participants times 5 if all participants provided the same word as their first response). According to the rank sum score of each picture, we obtained 174 words (16 pictures were given up due to the rank scores below 69). The scores varied between 70 and 162, with a mean of 92. We found in the 174 words there existed general consistency in terms of word-picture semantic relationship: 85 words were elicited from SS pictures, and 89 words from EE pictures.

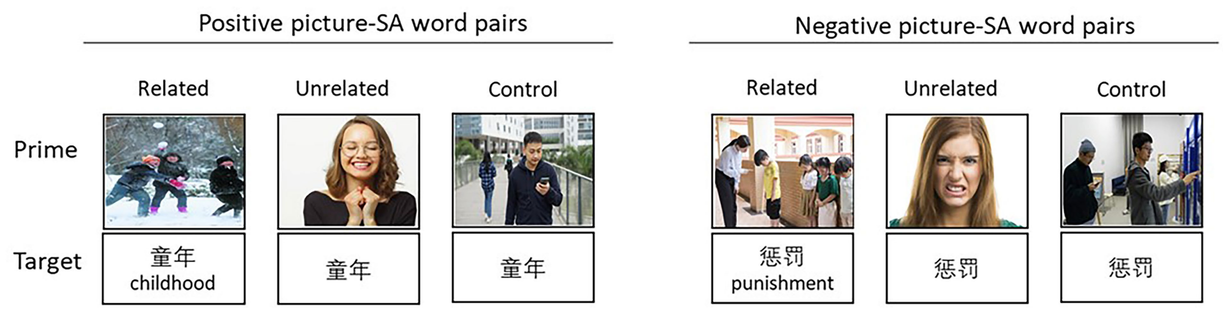

FIGURE 1
Examples of a social scene picture (left) and an emotional expression picture (right).

To insure a reliable social or emotional semantic relationship of a picture and its corresponding name, we additionally conducted a rating of semantic correlation of the 174 picture-word pairs.

The rating of social or emotional semantic correlation between each picture and its name

Participants

Sixty native Chinese speakers (12 males, mean age = 18.4, *SD* = 1.3) were divided randomly and equally into two groups. This rating study was conducted in a public psychology course of Xian Jiaotong University, and all participants read and signed consent to the study and received course credits for their participation.

Materials and procedure

One hundred ninety-four picture-word pairs were randomly divided into two parts (97 pairs in each), consisting of 174 pairs chosen from the previous step, and 20 unrelated pairs (without semantic and emotional association) used as filler stimuli to offset the fact that all picture-word pairs were related. The prime pictures in fillers were neutral pictures that were selected from CAPS (Bai et al., 2005).

Two group participants were asked to rate social or emotional semantic relationship of each picture-word pair on the Likert scale ranging from 1 to 7, with 1 indicating *extremely unrelated*, and 7 indicating *extremely related*. All participants completed the task within 15 min.

Results

Three participants were removed from the analysis due to a poor performance to filler stimuli (eight out of ten were wrong). We calculated the mean scores and standard deviations for each picture-word pair on semantic relationship in SPSS 26.0 and decided to consider picture-word pairs with a score of 6.13 or higher as the semantic related pairs. As a result, we chosen 72 SS picture-SA word pairs and 68 EE picture-EA word pairs.

In the next step, we collected subjective ratings for pictures and words on several important affective and/or lexical variables that were known to affect behavioural responses.

The ratings of affective and/or lexical variables for pictures and words

Participants

Fifty-two native Chinese speakers were recruited from Xian Jiaotong University, ranging in age from 18 to 25 years (29 males, mean age ± *SD* = 23.4 ± 1.3). They were randomly and equally assigned to complete either affective variables ratings for pictures or affective and lexical variables ratings for words. They received monetary compensation in the end for participation.

Materials and procedure

One group completed valence, arousal, abstractness, familiarity, and referent ratings for words on Likert scales ranging from 1 to 9 (1 meant extremely negative/calm/abstract/unfamiliar/label individual inner feelings, 9 meant extremely positive/arousing/concrete/familiar/derived from interpersonal interaction). The instructions for valence, arousal, abstractness, and familiarity referred to our prior work (Yao et al., 2017). The instruction for referent referred to the definitions of SA and EA concepts. Likewise, valence and arousal ratings for pictures were assessed by the other group. Appendix B presents the English translations of the instructions.

The rating task was implemented by the online questionnaire survey (SO JUMP) and was self-paced, completed in 1–3 sessions within 2 days.

Results

All participants completed the ratings tasks, and thus no participant's responses were removed. We calculated mean valence and arousal scores of each picture, as well as mean scores of each word on valence, arousal, abstractness, familiarity, and referent in SPSS 26.0.

Picture and its name (i.e., target word) that were presented to participants in the formal experiments were selected according to several criteria that were contrasted with one-way analysis of variance (ANOVA; see Table 1) and *post hoc* analyses with the Bonferroni correction ($\alpha < 0.05$): positive and negative SS/EE pictures were matched on arousal yet differed in valence. Positive and negative SA/EA target words were matched on arousal, abstractness, and familiarity, but differed in valence and referent. Descriptive statistics for selected pictures and words are summarized in Table 1.

As a result, the 48 social-semantic related pairs (24 positive and 24 negative SS picture-SA word pairs) and 48 emotional-semantic related pairs (24 positive and 24 negative EE picture-EA word pairs) were created. Part of picture-word pairs is presented in Appendix C.

Experiment 1: Effect of social experience on the recognition of SA concepts

In Experiment 1, we examined whether social experience that is provided by positive or negative SS pictures facilitates the responses to social-semantic related SA words in semantic priming. It was hypothesized that a significant semantic priming effect could be observed, with a quicker response to SA words that were preceded by SS vs. EE pictures (emotional-semantic related). Moreover, the social semantic priming effect of positive pairs might be different from negative pairs.

Methods

Participants

Thirty-four native Chinese speakers (18 males; 18–24 years old, mean age ± $SD = 19.8 \pm 2.1$) were all right-handed (Oldfield, 1971) with normal or corrected-to-normal vision. None of them had history of neurological or psychiatric disorders. Each participant signed a written informed consent before the experiment and received monetary compensation for their participation. The study was approved by the local Ethics Committee of Xian Jiaotong University.

Materials

Six experimental conditions were created according to target valence (positive, negative) and social-semantic relatedness (social related, social unrelated, control) of picture-SA word pairs, with 24 items in each condition (see Figure 2 for examples). Two conditions featured either positive or negative social semantic association between SS pictures and SA words, which were selected from the norming study. No social-semantic association was included in the other two conditions, in which SA words were preceded by EE pictures. All pairs shared either positive or negative valence but without any social semantic association. The remaining two conditions with neutral pictures as primes were used as control conditions, in which positive and negative SA words were preceded by neutral pictures. The 24 neutral pictures were selected from ISIEA database (the image database of social inclusion/exclusion in Asian young adults. Zheng et al., 2021) and described people in social situations that did not involve interpersonal interaction and affective meanings. The neutral pictures significantly differed in valence ($F_{2,117} = 649.7$, $p < 0.001$; 5.06 ± 0.24) and arousal ($F_{2,117} = 299.0$, $p < 0.001$; 4.21 ± 0.14) from positive (valence = 6.67 ± 0.43; arousal = 6.19 ± 0.39) and negative (valence = 3.02 ± 0.63; arousal = 6.26 ± 0.40) pictures.

The relatedness of picture-word pairs in social semantically unrelated and control conditions were rated by another sample of 18 participants using the 7-point scale (a higher score indicates a higher level of social-semantic association). The means and *SDs* of semantic relatedness ratings for all pairs are presented in Table 2.

TABLE 1 Descriptive statistics for selected pictures and words samples.

Variables			Valence	Arousal	Abstractness	Familiarity	Referent
Stimuli type							
Picture	Positive	SS	6.73 ± 0.44	6.20 ± 0.45			
		EE	6.61 ± 0.42	6.17 ± 0.34			
	Negative	SS	2.99 ± 0.66	6.29 ± 0.47			
		EE	3.05 ± 0.61	6.24 ± 0.33			
One-way ANOVA each factor			$F_{3,92} = 358.5$, $p < 0.001$	$F_{3,92} = 0.36$, $p = 0.78$, n.s.			
Word	Positive	SA	6.33 ± 0.46	6.09 ± 0.37	2.28 ± 0.52	6.26 ± 0.50	6.52 ± 0.65
		EA	6.44 ± 0.35	5.92 ± 0.51	2.43 ± 0.63	6.33 ± 0.42	2.14 ± 0.51
	Negative	SA	2.71 ± 0.76	6.02 ± 0.56	2.53 ± 0.60	6.02 ± 0.60	6.94 ± 0.49
		EA	2.87 ± 0.88	6.11 ± 0.54	2.48 ± 0.52	6.26 ± 0.63	1.87 ± 0.42
One-way ANOVA each factor			$F_{3,92} = 246.1$, $p < 0.001$	$F_{3,92} = 0.68$, $p = 0.57$, n.s.	$F_{3,92} = 0.86$, $p = 0.47$, n.s.	$F_{3,92} = 1.44$, $p = 0.24$, n.s.	$F_{3,92} = 647.0$, $p < 0.001$

Means of Valence (1, Negative to 9, Positive), Arousal (1, Calming to 9, Arousing), Abstractness (1, Abstract to 9, Concrete), Familiarity (1, Unfamiliar to 9, Familiarity), and Referent (1, Inner feelings to 9, Human interaction). n.s., nonsignificant; ANOVA, analysis of variance; SS, Social Scene pictures; EE, Emotional Expression pictures; SA, Social Abstract words; EA, Emotional Abstract words.

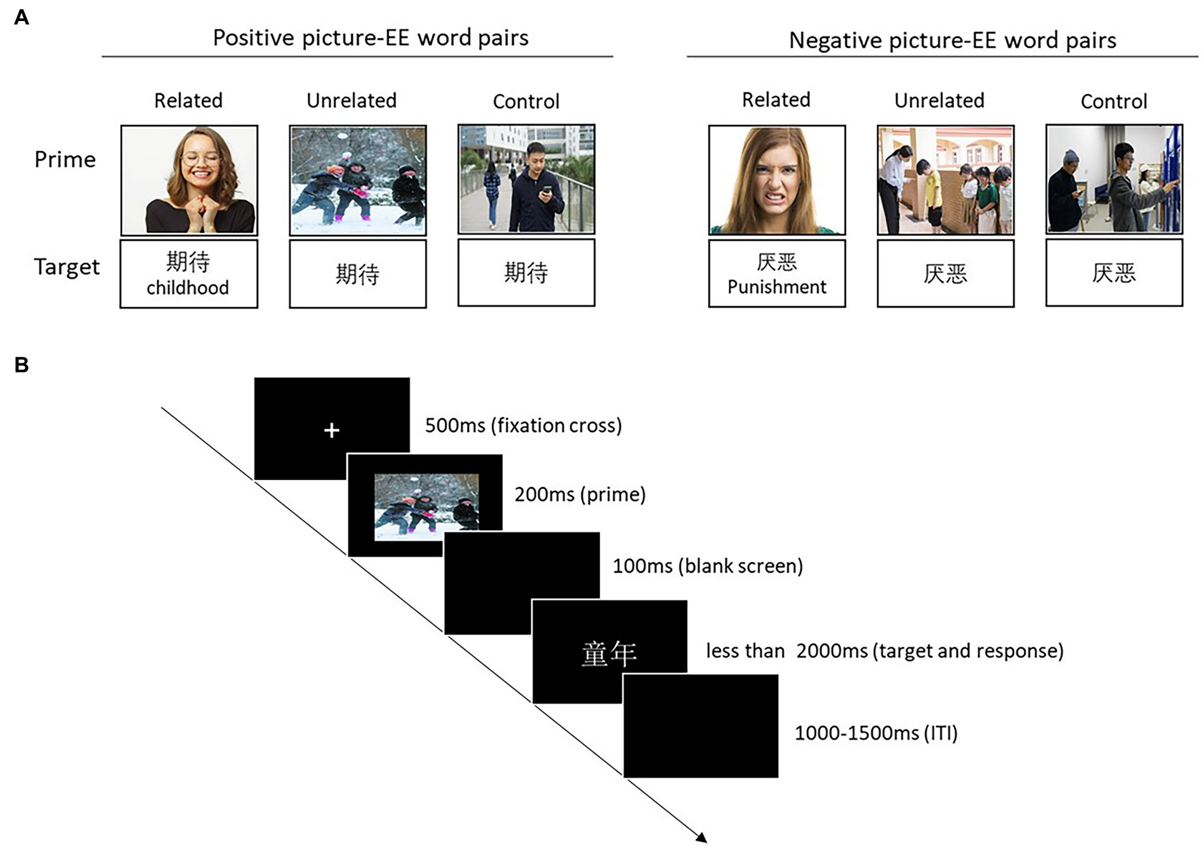

FIGURE 2
(A) Six experimental conditions of Experiment 1. **(B)** Six experimental conditions of Experiment 2.

TABLE 2 The rating scores of semantic relationships between prime pictures and SA words.

	SA words paired with different primes	Semantic relationship
Related	positive SS pictures	6.28 ± 0.98
	negative SS pictures	6.19 ± 1.07
Unrelated	positive EE pictures	2.11 ± 1.68
	negative EE pictures	1.94 ± 0.87
Control	neutral pictures-positive SA words	2.01 ± 1.05
	neutral pictures-negative SA words	1.92 ± 1.25

SS pictures, Social Scene pictures; EE pictures, Emotional Expression pictures; SA words, Social Abstract words; EA words, Emotional Abstract words.

In addition, we used 144 picture-pseudoword pairs. The pseudowords, all pronounceable, were generated by altering one random character within different real words. In short, there were 264 experimental pairs, consisting of 96 social-semantic related or social-semantic unrelated picture-SA word pairs that had either positive or negative valence (24 pairs in four conditions), 24 neutral picture-SA word pairs, and 144 picture-pseudoword pairs.

Task and procedure

Participants performed the lexical decision task in separate sound-proof booths. They were told that a picture would be briefly presented on the screen and be immediately followed by a word. They were indicated to respond as quickly and accurately as possible whether the word was a real word or a pseudoword by pressing the "Z" and "M" keys on the keyboard (assignment of the two keys to response categories was counterbalanced across participants).

All 288 trials (24 neutral picture-SA word pairs were repeated one time) were shown in four blocks of 72 trials each. Two of the blocks had positive SA words as targets, with each containing 12 social-semantic related, 12 social-semantic unrelated, 12 neutral, and 36 pseudoword pairs. The other two blocks had negative SA words as targets, and each block contained the same proportion of pairs as positive SA blocks. The order of pair presentation in each block and the order of the blocks was randomized for each participant.

Stimuli (prime pictures: 400 × 245 pixels; target words: Song typeface, size: 36) and instructions were presented in white letters over black background on a 21-in. monitor. Each trial started with the presentation of a fixation cross for 500 ms, followed by a prime picture for 200 ms. After the prime, a blank screen was shown for

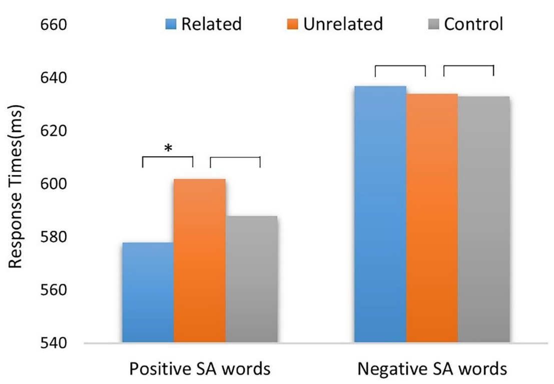

FIGURE 3
Trial procedure for Experiment 1. "童(tong2)年(nian2) means "childhood."

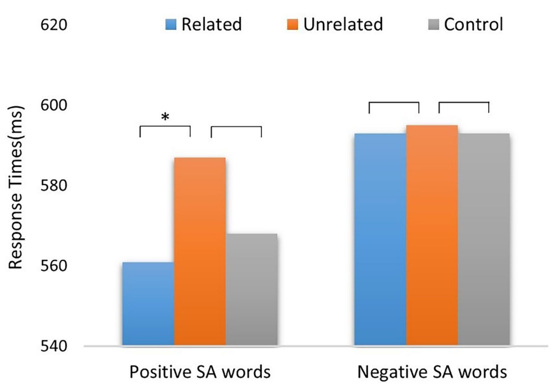

FIGURE 4
Response times of positive/negative SA words in social-semantic related, unrelated, and control conditions. SA words=Social Abstract words.

100 ms before the target was presented until the participant responded or 2,000 ms elapsed. The inter-trial interval was 1,000–1,500 ms (see Figure 3). Prior to the experiment trials, each participant performed 12 practice trials (these trials did not appear in the formal experiment) to prove that they had completely understood the procedure and correct key presses. All experiments were programmed using E-Prime 3.0 (Psychology Software Tools Inc., Sharpsburg, PA).

Results

One participant with an overall accuracy below 50% was removed from the analyses. In the remaining 33 participants, overall accuracy was high (98.8%) and did not differ between experimental conditions (range: 97.5–99.6%). Therefore, consequent analysis concentrated on response times (RTs). We excluded from the analyses mean RTs above or below 2.5 standard deviations from the mean, and only analyzed RTs for correct responses to target stimuli.

A repeated-measures ANOVA was run on RTs in the six prime-target conditions: 2 (target valence: positive vs. negative) × 3 (social-semantic relationship: related vs. unrelated vs. control). The results revealed a significant main effect of target valence ($F_{1,32}=34.5$, $p=0.001$, $\eta_p^2=0.52$), with slower responses to negative SA words (634.2 ± 13.3 ms) compared to positive SA words (589.1 ± 13.1 ms). A main effect of social-semantic relationship of picture-word pairs was significant ($F_{2,64}=3.72$, $p=0.03$, $\eta_p^2=0.80$), with responses to social-semantic unrelated pairs (617.7 ± 12.9 ms) being slower than to social-semantic related (607.1 ± 12.2 ms) and control pairs (610.1 ± 13.5 ms). A significant interaction was found between target valence and social-semantic relationship ($F_{2,64}=3.95$, $p=0.03$, $\eta_p^2=0.11$; Figure 4). The simple-effect analysis showed that responses to positive SA words were significantly faster in semantically related (577.8 ± 12.2 ms) condition than in unrelated (601.9 ± 14.0 ms) and control (587.7 ± 14.4 ms) conditions ($F_{2,64}=7.92$, $p<0.001$). However, for negative SA words, no significant differences between related (636.5 ± 13.8 ms), unrelated (633.5 ± 13.3 ms), and control (632.5 ± 14.4 ms) conditions were observed ($F_{2,64}=0.2$, $p=0.82$).

Discussion

In Experiment 1, positive SA words were facilitated by the corresponding positive SS pictures compared with positive EE pictures, showing a significant social-semantic priming effect. However, for negative SS picture-SA words, no significant difference in RTs was found between related and unrelated conditions. The pattern of results suggests that positive SA words could more readily benefit from social experiential information of prime pictures to strengthen the semantic association between them. A possible reason for this result is that most people usually live in a normal social environment, in which people tend to watch, hear, and experience a healthy and positive social interaction. Similar positive bias during processing of words was also observed in many previous studies (Kanske and Kotz, 2007; Hinojosa and Me, 2010; Bayer et al., 2012), which was explained by that the human brain is more reactive to the valence of positive relative to negative words (Yang et al., 2013).

Experiment 2: Effect of social experience on the recognition of EA words

In Experiment 2, EA words with positive or negative were used as targets, which were preceded by affectively consistent EE or SS prime pictures. We expected that, if EA words are more detached form social experience than SA concepts, then EA target words would not be facilitated by the SS pictures instead of the

TABLE 3 The rating scores of semantic relationships between prime pictures and EA words.

	EA words paired with different primes	Semantic relationship
Related	positive EE pictures	6.41 ± 1.36
	negative EE pictures	6.22 ± 1.17
Unrelated	positive SS pictures	2.18 ± 1.58
	negative SS pictures	2.11 ± 1.07
Baseline	neutral pictures-positive EA words	1.99 ± 1.10
	neutral pictures-negative EA words	2.02 ± 0.91

EE pictures, Emotional Expression pictures; SS pictures, Social Scene pictures; EA words; Emotional Abstract words; SA words, Social Abstract words.

FIGURE 5
Response times of positive/negative EA target words in semantically related, unrelated, and control conditions. EA words=Emotional Abstract words.

corresponding EE pictures, showing a different pattern from SA words.

Methods

Participants

Thirty-nine university students (25 males; 17–23 years old, mean age ± $SD = 22.3 \pm 2.4$) participated in Experiment 2 and received financial compensation for participation (see Experiment 1 for further details). They all gave a written informed consent prior to the experiment.

Materials

Similar to the proportion of Materials in Experiment 1, there were also 264 experimental pairs, with the exception of using EA words as targets (see Figure 2B for examples). Specifically, the stimulus set included 48 emotional-semantic related EE picture-EA word pairs (24 positive, 24 negative), 48 emotional-semantic unrelated SS picture-EA word pairs (24 positive, 24 negative), 24 neutral picture-EA word pairs, and 144 picture-pseudoword pairs. The semantic relatedness of all picture-EA word pairs was also rated using the 7-point scale, with the same sample as in Experiment 1. The rating scores are shown in Table 3.

Task and procedure

The experimental task and procedure were the same as in Experiment 1.

Results

We excluded from the analyses mean response times (RTs) above or below 2.5 standard deviations from the mean, and only analyzed RTs for correct responses to target stimuli, because the accuracy for each trial in all conditions was high (98.2%, range: 97.2–99.4%) and did not differ across conditions.

A repeated-measure AVOVA on RTs in six experimental conditions and revealed a significant main effect of target valence ($F_{1,38} = 17.7$, $p < 0.001$, $\eta_p^2 = 0.32$), with longer RTs in negative pairs (593.6 ± 11.0 ms) than in positive pairs (571.8 ± 10.6 ms). A main effect of emotional-semantic relationship was significant ($F_{2,76} = 6.52$, $p = 0.002$, $\eta_p^2 = 0.15$): responses to related (577.2 ± 10.8 ms) and control conditions (580.1 ± 10.6 ms) were faster than responses to unrelated conditions (590.8 ± 10.8 ms). A significant interaction between target valence and semantic relationship was found ($F_{2,76} = 4.11$, $p = 0.02$, $\eta_p^2 = 0.1$; Figure 5). The simple-effect analysis showed that RTs to positive EA words were significantly longer in semantically unrelated (586.5 ± 11.4 ms) conditions compared with related (561.4 ± 10.8 ms) and control (567.5 ± 10.8 ms) conditions ($F_{2,76} = 12.45$, $p < 0.001$), but no significant differences in RTs to negative EA words whether in related (593.0 ± 11.6 ms), unrelated (595.1 ± 11.5 ms), and control (592.8 ± 11.7 ms) conditions ($F_{2,76} = 0.09$, $p = 0.92$).

Discussion

The results of Experiment 2 shown that only positive EA words were facilitated by emotional-semantic related positive EE pictures relative to positive SS pictures. That is, a significant emotional-semantic priming effect was found for positive EA words, but not for negative EA words. We inferred that positive EA words describing inner positive feelings might more readily benefit from our facial displays and body gestures, thus evoking a tighter emotional-semantic association between positive EE pictures and EA words. By comparison, social experiential knowledge that underlies positive SS pictures seems not to offer an additional resource to accelerate the recognition of positive EA words.

Similar to negative SA words in Experiment 1, negative EA words were not influenced, whether they were preceded by negative EE or SS pictures. This result is consistent with previous studies (Rossell and Nobre, 2004; Sass et al., 2012), suggesting that primes' negative information inhibits the spread of activation between related concepts, so that target words cannot extract related emotional-semantic content from pictures to build a closer semantic relatedness, thus showing a null effect for negative primes.

General discussion

In the present study, a picture-word semantic priming paradigm was designed to explore the distinct role of social experience in the grounding of SA and EA concepts. The SA and EA words that shared similar affective and lexical variables were used as targets in Experiments 1 and 2, respectively. The results of the two experiments shown that semantic priming effects were observed for both positive picture-SA word pairs and positive picture-EA word pairs, with quicker responses to semantically related pairs than to semantically unrelated pairs. Note that positive SA words were facilitated by the corresponding SS pictures, whereas positive EA words were facilitated by the corresponding EE pictures instead of SS pictures. This pattern of results suggests that positive social experience from real-life scenes (i.e., SS pictures) facilitates the recognition of related SA words, but not of positive EA words. Moreover, such facilitation was not observed in negative picture-SA/EA word conditions. Overall, these findings confirm the WAT view, emphasizing a crucial role of social experience for abstract concepts, and further reveal that social experience could be an embodied dimension for specifically characterizing SA concepts and distinguishing SA concepts from other types of abstract concepts, such as EA concepts, at least in the positive semantic priming context.

Our experiments show significant semantic priming effects in positive picture-SA/EA words, which are consistent with other studies exploring automatic semantic priming in a lexical decision task (e.g., Rossell and Nobre, 2004; Sass et al., 2012). Such priming effects can be explained by spreading activation within semantic networks. According to spreading activation theory (Neely, 1991), activation is considered to spread from a prime to a target if the two share a closer semantic association in semantic memory, thereby influencing decisions to targets. Thus, quicker responses to positive SA words in related vs. unrelated conditions can be explained by the fact that positive SA words more readily benefit from social experience conveyed by the related SS pictures, and thus strengthen the semantic association between the two. By comparison, positive EA words are more easily accelerated by facial expressions and body gestures that are provided by the corresponding EE pictures. In other words, social experience that is provided by positive SA pictures did not facilitate the recognition of positive EA words. Such findings are in accordance with our hypothesis, showing that sub-kinds of abstract concepts may not equally benefit from social experience, at least in positive semantic priming context.

Moreover, our findings confirm the WAT, which emphasizes the relationship between social experience and abstract concepts (Borghi et al., 2017, 2019; Pecher, 2018; Davis et al., 2020), and also support recent empirical works reporting a facilitating role of sociality in the processing of abstract concepts (Mellem et al., 2016; McRae et al., 2018; Zdrazilova et al., 2018; Fini et al., 2021). For example, McRae et al. (2018) reported that pictures depicting real-world social scenes (e.g., *two girls sharing a corn cob*) could facilitate the processing of related abstract words (e.g., *friendship*) in a lexical decision task, and vice versa. In our study, positive SA and EA words shared similar affective and lexical variables in a parallel semantic priming context, and thus the difference between the two in the recognition performance could be ascribed to the fact that they are grounded in varied degrees of social experience. Specifically, positive SA words are characterized by more social features compared to positive EA words. This finding provides empirical evidence to show that abstract concepts may be quite different from one another (Connell et al., 2018; Harpaintner et al., 2018; Villani et al., 2019; Wang and Bi, 2019; Zhang et al., 2019), and also supports recent studies that claimed abstract concepts should be studied using a category-specific approach (Ghio et al., 2016; Desai et al., 2018; Mkrtychian et al., 2019). In this sense, social experience is expected to be an important embodied dimension to characterize SA concepts and distinguish them from other different categories of abstract concepts.

However, the results of the two experiments consistently indicated that no significant differences in response times were observed between semantically related and unrelated negative picture-SA/EA words pairs, suggesting that the automatic spreading of activation did not occur between negative primes and targets. This is probably due to the negativity of the information encoded in the primes, which inhibits such spreading activation. As expected, the effect of social experience on abstract concepts was modulated by their valence, which is in line with previous findings with regard to the different semantic priming between positive and negative primes (Sass et al., 2012). For example, Rossell and Nobre (2004) analyzed the effect of valence on semantic priming and observed a significant priming effect for positive stimuli, a null effect for fearful stimuli, and an inhibited effect for sad stimuli. According to the spreading inhibition hypothesis (Clore and Storbeck, 2006), one possible reason for the null effects for negative picture-word pairs in Experiment 1 and 2 is that there exists a different organization of positive and negative information in semantic memory. Positive information of primes can increase the accessibility of prime-target pairs, by contrast, negative information of primes inhibits it, which makes the spreading of activation between connected nodes more difficult. Additionally, some researchers clearly proposed that people seem to perceive positive information as more compatible with negative information (Unkelbacha et al., 2020), because the human brain is more reactive to the valence of positive relative to negative words (Yang et al., 2013).

Conclusion

In summary, our study demonstrated that social experience exerted a different role in the recognition of SA and EA concepts in a lexical priming-decision task. Specifically, the recognition of SA concepts could benefit from semantically related SS pictures in positive priming context, whereas EA concepts did not. This finding suggests that positive SA and EA concepts are grounded in different degrees of social experience. Thus, as a newly emerging embodied dimension, social experience may be capable of characterizing the key features of SA concepts, thus effectively distinguishing them from different kinds of abstract concepts.

Although it is increasingly apparent that social experience is a constitutive part of abstract concepts, work in this area is still in its early stages. Our study provides preliminary evidence to support the varied facilitating effect of social experience on different sub-categories of abstract concepts, and such effect is limited to positive SA concepts in semantic priming. One important limitation of the present study is that only a lexical decision-priming task and behavioral measure were employed, which may limit the application of our findings to the studies of the relationship between abstract concepts and sociality. Moreover, apart from social scene pictures, social situational sentences or social experience also provide specific social experiential information. Therefore, further research should extend this topic to other cognition tasks and materials, and further explore the role of social experience in the grounding of different kinds of abstract concepts.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Xi'an Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ZY designed the study, analyzed the data, wrote the paper, and approved the final version and is accountable for all aspects of the work. FW were involved in data analysis work and corrected grammar. YC, PY, and RZ helped to collect and interpret the data. All authors contributed to the article and approved the submitted version.

Acknowledgments

This research was supported by grants from the Social Science Foundation of Shaanxi Province (2021K013) and the "Young Talent Support Plan" of Xi'an Jiaotong University to ZY.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.912176/full#supplementary-material

References

Bai, L., Ma, H., Huang, Y. X., and Luo, Y. J. (2005). The development of native Chinese affective picture system–a pretest in 46 college students. *Chin. Ment. Health J.* 19, 719–722. doi: 10.3321/j.issn:1000-6729.2005.11.001

Barsalou, L. W., Santos, A., Simmons, K. W., and Wilson, C. D. (2008). "Language and simulations in conceptual processing," in *Symbols, Embodiment and Meaning.* (eds.) M. De Vega , A. M. Glenberg and A. C. Graesser (Oxford, UK: Oxford University Press), 245–284.

Bayer, M., Sommer, W., and Schacht, A. (2012). P1 and beyond: functional separation of multiple emotion effects in word recognition. *Psychophysiology* 49, 959–969. doi: 10.1111/j.1469-8986.2012.01381.x

Bechtold, L., Bellebaum, C., Egan, S., Tettamanti, M., and Ghio, M. (2019). The role of experience for abstract concepts: expertise modulates the electrophysiological correlates of mathematical word processing. *Brain Lang.* 188, 1–10. doi: 10.1016/j.bandl.2018.10.002

Bolognesi, M., and Steen, G. (2018). Editors' introduction: abstract concepts: structure, processing, and Modeling. *Top. Cogniv. Sci.* 10, 490–500. doi: 10.1111/tops.12354

Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., and Tummolini, L. (2019). Words as social tools: language, sociality and inner grounding in abstract concepts. *Phys Life Rev* 29, 120–153. doi: 10.1016/j.plrev.2018.12.001

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., and Tummolini, L. (2017). The challenge of abstract concepts. *Psychol. Bull.* 143, 263–292. doi: 10.1037/bul0000089

Clore, G. L., and Storbeck, J., (2006). *Affect as Information About Liking, Efficacy, and Importance*. Psychology Press, New York, pp. 123–142.

Connell, L., Lynott, D., and Banks, B. (2018). Interoception: The forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosop. Trans. Royal Soc. B* 373:143. doi: 10.1098/rstb.2017.0143

Davis, C. P., Altmann, G. T., and Yee, E. (2020). Situational systematicity: A role for schema in understanding the differences between abstract and concrete concepts. *Cogn. Neuropsychol.* 37, 142–153. doi: 10.1080/02643294.2019.1710124

Desai, R. H., Reilly, M., and van Dam, W. (2018). The multifaceted abstract brain. *Philosop. Trans. Royal Soc. B* 373, 20170122. doi: 10.1098/rstb.2017.0122

Dreyer, F. R., and Pulvermüller, F. (2018). Abstract semantics in the motor system? An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex* 100, 52–70. doi: 10.1016/j.cortex.2017.10.021

Fini, C., Era, V., Da Rold, F., Candidi, M., and Borghi, A. M. (2021). Abstract concepts in interaction: the need of others when guessing abstract concepts smooths dyadic motor interactions. *R. Soc. Open Sci.* 8:201205. doi: 10.1098/rsos.201205

Ghio, M., Vaghi, M. M. S., Perani, D., and Tettamanti, M. (2016). Decoding the neural representation of fine-grained conceptual categories. *NeuroImage* 132, 93–103. doi: 10.1016/j.neuroimage.2016.02.009

Giffard, B., and Pratique, E. (2015). An exploration of the semantic network in Alzheimer' s disease: influence of emotion and concreteness of concepts science direct An exploration of the semantic network and concreteness of concepts. *Cortex* 69, 201–211. doi: 10.1016/j.cortex.2015.05.020

Glaser, W. R. (1992). Picture naming. *Cognition* 42, 61–105. doi: 10.1016/0010-0277(92)90040-O

Harpaintner, M., Trumpp, N. M., and Kiefer, M. (2018). The semantic content of abstract concepts: A property listing study of 296 abstract words. *Front. Psychol.* 9:1748. doi: 10.3389/fpsyg.2018.01748

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., et al. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychol. Bull.* 139, 1062–1089. doi: 10.1037/a0031309

Hinojosa, A., and Me, C. (2010). Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology* 47, 748–757. doi: 10.1111/j.1469-8986.2010.00982.x

Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044

Kazanas, S. A., and Altarriba, J. (2015). Emotion word type and affective valence priming at a long stimulus onset asynchrony. *Lang. Speech* 59, 339–352. doi: 10.1177/0023830915590677

Kiefer, M., and Harpaintner, M. (2020). Varieties of abstract concepts and their grounding in perception or action. *Open Psychol.* 2, 119–137. doi: 10.1515/psych-2020-0104

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, P. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. General* 140, 14–34. doi: 10.1037/a0021446

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective picture system (IAPS): Affective Ratings of Pictures and Instruction Manual*. Gainesville, FL: Center for the Study of Emotion & Attention.

Louwerse, M. M., and Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition* 114, 96–104. doi: 10.1016/j.cognition.2009.09.002

McRae, K., Nedjadrasul, D., Pau, R., Lo, B. P. H., and King, L. (2018). Abstract concepts and pictures of real-world situations activate one Another. *Top. Cogn. Sci.* 10, 518–532. doi: 10.1111/tops.12328

Mellem, M. S., Jasmin, K. M., Peng, C., and Martin, A. (2016). Sentence processing in anterior superior temporal cortex shows a social-emotional bias. *Neuropsychologia* 89, 217–224. doi: 10.1016/j.neuropsychologia.2016.06.019

Meyer, D., and Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *J. Exp. Psychol.* 90, 227–234. doi: 10.1037/h0031564

Mkrtychian, N., Blagovechtchenski, E., Kurmakaeva, D., Gnedykh, D., Kostromina, S., and Shtyrov, Y. (2019). Concrete vs. abstract semantics: From mental representations to functional brain mapping. *Front. Hum. Neurosci.* 13:267. doi: 10.3389/fnhum.2019.00267

Moseley, R., Carota, F., Hauk, O., Mohr, B., and Pulvermüller, F. (2012). A role for the motor system in binding abstract emotional meaning. *Cereb. Cortex* 22, 1634–1647. doi: 10.1093/cercor/bhr238

Neely, J. H. (1991). "Semantic priming effects in visual word recognition: A selective review of current findings and theories," in *Basic Processes in reading: Visual word Recognition*. eds. D. Besner and G. W. Humphreys (Hillsdale, NJ: Erlbaum), 264–336.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pecher, D. (2018). Curb your embodiment. *Top. Cogn. Sci.* 10, 501–517. doi: 10.1111/tops.12311

Recchia, G., and Jones, M. N. (2012). The semantic richness of abstract concepts. *Front. Hum. Neurosci.* 6:315. doi: 10.3389/fnhum.2012.00315

Rossell, S. L., and Nobre, A. C. (2004). Semantic priming of different affective categories. *Emotion* 4, 354–363. doi: 10.1037/1528-3542.4.4.354

Roversi, C., Borghi, A. M., and Tummolini, L. (2013). A marriage is an artefact and not a walk that we take together: An experimental study on the categorization of artefacts. *Rev. Philos. Psychol.* 4, 527–542. doi: 10.1007/s13164-013-0150-7

Sass, K., Habel, U., Sachs, O., Huber, W., Gauggel, S., and Kircher, T. (2012). The influence of emotional associations on the neural correlates of semantic priming. *Hum. Brain Mapp.* 33, 676–694. doi: 10.1002/hbm.21241

Unkelbacha, C., Alvesa, H., and Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: explaining the differential processing of positive and negative information. *Adv. Exp. Soc. Psychol.* 62, 115–187. doi: 10.1016/bs.aesp.2020.04.005

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025

Villani, C., Lugli, L., Liuzza, M., and Borghi, A. (2019). Varieties of abstract concepts and their multiple dimensions. *Lang. Cogn.* 11, 403–430. doi: 10.1017/langcog.2019.23

Wang, X. S., and Bi, Y. C. (2019). The cognitive and neural bases of abstract concepts. *Acta Physiologica Sinica* 25, 117–126. doi: 10.13294/j.aps.2018.0078

Wiemer-Hastings, K., and Xu, X. (2005). Content differences for abstract and concrete concepts. *Cogn. Sci.* 29, 719–736. doi: 10.1207/s15516709cog0000_33

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., and Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia* 49, 1105–1127. doi: 10.1016/j.neuropsychologia.2010.12.032

Yang, J., Zeng, J., Meng, X., Zhu, L., Yuan, J., Li, H., et al. (2013). Positive words or negative words: whose valence strength are we more sensitive to? *Brain Res.* 1533, 91–104. doi: 10.1016/j.brainres.2013.08.020

Yao, Z., Wang, Y., Lu, B., and Zhu, X. (2019). Effects of valence and arousal on affective priming vary with the degree of affective experience denoted by words. *Int. J. Psychophysiol.* 140, 15–25. doi: 10.1016/j.ijpsycho.2019.03.011

Yao, Z., Wu, J., Zhang, Y., and Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behav. Res. Methods* 49, 1374–1385. doi: 10.3758/s13428-016-0793-2

Yee, E. (2019). Abstraction and concepts: when, how, where, what and why? *Lang. Cogn. Neurosci.* 34, 1257–1265. doi: 10.1080/23273798.2019.1660797

Zdrazilova, L., Sidhu, D. M., and Pexman, P. M. (2018). Communicating abstract meaning: concepts revealed in words and gestures. *Philosop. Trans. Royal Soc. B* 373, 20170138. doi: 10.1098/rstb.2017.0138

Zhang, J., Wu, C., Yuan, Y., and Meng, Y. (2019). Differentiating emotion-label words and emotion-laden words in emotion conflict: an ERP study. *Experi. Brain Res.* 237, 2423–2430. doi: 10.1007/s00221-019-05600-4

Zheng, Z., Li, S., Mo, L., Chen, W., and Zhang, D. (2021). ISIEA: An image database of social inclusion and exclusion in young Asian adults. *Behav. Res. Methods* 12:1735. doi: 10.3758/s13428-021-01736-w

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fpsyg.2022.906154

OPEN ACCESS

EDITED BY
Christoph Scheepers,
University of Glasgow, United Kingdom

REVIEWED BY
Shari R. Baum,
McGill University, Canada
Julia Elisabeth Hofweber,
University College London,
United Kingdom

*CORRESPONDENCE
Huili Wang
huiliw1966@outlook.com

SPECIALTY SECTION
This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

RECEIVED 28 March 2022
ACCEPTED 11 July 2022
PUBLISHED 06 September 2022

CITATION
Zang A, de Vega M, Fu Y, Wang H and
Beltrán D (2022) Language switching
may facilitate the processing
of negative responses.
Front. Psychol. 13:906154.
doi: 10.3389/fpsyg.2022.906154

Language switching may facilitate the processing of negative responses

Anqi Zang[1], Manuel de Vega[1], Yang Fu[1], Huili Wang[2]* and David Beltrán[1,3]

[1]Instituto Universitario de Neurociencia, Universidad de La Laguna, Santa Cruz de Tenerife, Spain, [2]School of Foreign Languages, Zhejiang University City College, Hangzhou, China, [3]Departamento de Psicología Básica I, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

It has been proposed that processing sentential negation recruits the neural network of inhibitory control (de Vega et al., 2016; Beltrán et al., 2021). In addition, inhibition mechanisms also play a role in switching languages for bilinguals (Kroll et al., 2015). Since both processes may share inhibitory resources, the current study explored for the first time whether and how language-switching influences the processing of negation. To this end, two groups of Spanish-English bilinguals participated in an encoding-verification memory task. They read short stories involving the same two protagonists (Montse and Jordi), referring to their activities in four different scenarios in Spanish or English. Following each story, the participants received verification questions requiring "yes" or "no" responses depending on whether a given fact was correctly referred to one of the protagonists. Some of the verification questions were in the story's original language (non-switch condition) and others in the alternate language (switch condition). Results revealed that language-switching facilitated negative responses compared to affirmative responses, exclusively for questions switching from dominant language (L1) to non-dominant language (L2). This effect might reflect that the domain-general mechanisms of inhibitory control are recruited at least partially for both language switch and negation process simultaneously, although this phenomenon is modulated by language dominance.

KEYWORDS

language switching, negation processing, inhibitory mechanism, cognitive control, bilinguals

Introduction

Imagine you are a Spanish-English bilingual, you read a news item from a Spanish newspaper and then you share the news item with two friends, one a Spanish monolingual and the other an English monolingual. After telling them the news, they began to ask you many details in Spanish or English, such as "Who had the accident? Mary?" and you replied "No," or "¿Quién conducía el coche? ¿Peter?" [Who was driving the car? Peter?] and you answer "Sí" [Yes]. This is an example of how learning and retrieving linguistic information can rely on the same language (L1 – L1), or on a different language (L1 – L2) to produce an affirmative or negative answer in the appropriate language. In this situation, you retrieve information in the original language, but, to answer some questions, you eventually need to switch to another language. To do that, one remarkable ability of bilinguals is cognitive control or monitoring to discern between their two languages and select one of them while suppressing the other to minimize the interference. Numerous studies have investigated how language control plays a crucial role in bilinguals (Green, 1998; Meuter and Allport, 1999; Abutalebi and Green, 2008), giving prominence to the demands for inhibitory resources. Yet, cognitive control and inhibition are also important for many aspects of language processing, including comprehension and production of negation (de Vega et al., 2016; Beltrán et al., 2019; Liu B. et al., 2020; Dudschig et al., 2021; Vitale et al., 2021). Therefore, the bilingual scenario above raises some questions: Do language switching and negative responses share the same inhibitory resources? And if so, how does the former influence the latter?

The role of cognitive control in bilinguals has been studied with the so-called language-switching paradigm. This paradigm requires the naming of pictures or digits in either of two languages, depending on explicit cues (Meuter and Allport, 1999), a pre-ordered sequence (Declerck et al., 2013), or a voluntary selection (Liu H. et al., 2020). The typical finding is that switching from one language to another requires a longer response time than keeping the same language (Costa and Santesteban, 2004; Declerck et al., 2013; Declerck and Philipp, 2015b). Many researchers have reported an asymmetric switching cost effect, with larger switching costs to shift from less dominant (L2) to dominant language (L1) than to shift from dominant (L1) to less dominant language (L2), which has been attributed to differential demands on inhibitory control (De Bot, 1992; Green, 1998; Meuter and Allport, 1999). That is, producing in a specific language involves inhibition of the non-target language, but more inhibition is needed to suppress the irrelevant L1 when producing L2 (see a review, Gade et al., 2021).

Most of the research in this field has focused on bilinguals' production of isolated words. However, some research also reported language switching effects on comprehension (Bultena et al., 2015; Wang, 2015) and long-term memory (Marian and Neisser, 2000; Matsumoto and Stanny, 2006; Marian and Kaushanskaya, 2007). In general, according to the encoding specificity principle, matching features of encoding and retrieval contexts facilitates recall, in comparison with mismatching encoding and retrieval contexts (see Tulving and Thomson, 1973; Davies and Thomson, 1988; for a review). Applying this principle to the field of bilingualism, several studies have examined language-dependent effects on memory, typically reporting that recall is better when the language of retrieval matched the language of encoding rather than when they are mismatched. This language matching advantage occurs in autobiographical memory (Marian and Neisser, 2000; Matsumoto and Stanny, 2006), semantic memory for world knowledge (Marian and Kaushanskaya, 2007), academic-type memory (Marian and Fausey, 2006), and narrative stories (Wang, 2022).

Negation processing has often been associated with the suppression of negated information (see Beltrán et al., 2021 for a recent overview). Indeed, several lines of research support the hypothesis that negation has inhibition-like effects. One approach reveals that negation modulates embodied effects during the comprehension of action language, as in the case of the reduction of motor interference effects in behavioral studies (Aravena et al., 2012; Bartoli et al., 2013; García-Marco et al., 2019), and the reduction of activation of the motor and premotor cortex reported by neuroimaging studies with fMRI technique (Tettamanti et al., 2008; Tomasino et al., 2010). Similarly, non-invasive brain stimulation studies identified a larger cortical silence period (a measure of inhibition in the GABAergic system) associated with negation when single-pulse TMS (Transcranial Magnetic Stimulation), which was used for stimulating peripheral nerves with a similar mechanism of activation as for electrical stimulation (Terao and Ugawa, 2002), was applied to M1 during the comprehension of action verbs (Papeo et al., 2016). Another approach uses the probe recognition paradigm to assess the activation level and the recall performance for negated concepts compared to affirmed concepts. The typical results showed longer latencies and higher error rate for negated concepts compared to non-negated concepts, indicating less accessibility for negated concepts, probably because negation interferes with (or inhibits) conventional concept encoding in working memory (e.g., MacDonald and Just, 1989; Kaup and Zwaan, 2003; Mayo et al., 2004; Orenes et al., 2014). Finally, EEG studies have demonstrated that negation recruits mechanisms of inhibitory control (de Vega et al., 2016; Beltrán et al., 2018, 2019; Liu B. et al., 2020). For instance, de Vega et al. (2016) provided the first evidence that understanding negative action sentences interacts with the processes required to suppress a prominent motor response in a concurrent Go/NoGo task, modulating the frontal theta rhythm, which is considered a typical marker of response inhibition.

The above studies mainly focus on the processing of sentential negation, that is, how sentences with a negative marker are understood. Yet, people produce negations as much as understand them. Thus, developmental studies have shown that children begin to use negative responses (no/not) very early, during the second year of life, to reject an object, or to stop or prevent an imminent action, establishing thus a strong association between the verbal markers of negation and the rejection and prevention of an action. Moreover, in this early stage of linguistic development, the child often use negation for self-prohibition, when she is about to engage in a forbidden action (e.g., Bloom, 1970; Pea, 1980; Choi, 1988). In fact, we can assume that inhibitory control underlies production of negations since the early childhood.

One important pragmatic function of negations is denial (Bloom, 1970). Denial occurs when a negative utterance is produced in response to a question that refers to a false content; for example, responding "no" when asked "Is this work written in Spanish?" Interestingly, verification tasks involving affirmative or negative responses (denials) have been widely used in studies of language and memory. Typically, participants receive statements referred to semantic memory contents ("Do cats eat vegetables?"), world knowledge ("Has Donald Trump been president?"), pictures or episodic memories about previously learned content and they simply have to answer yes or no. In a pioneering study, Craik and Tulving (1975) utilized a memory retrieval paradigm to identify distinct effects of response polarity on memory in their study depending on levels of processing. Participants had to initially encode words at various levels of processing, such as whether they were written in capital letters (shallow encoding) or whether they fit into a semantic category or sentence structure (deep encoding). In a posterior incidental memory test, they found that negative (no) responses had poorer recall than affirmative (yes) ones, particularly under deep encoding circumstances, supporting the hypothesis that negation might induce forgetting by weakening encoding strength. A few behavioral investigations have shown that the impact of negation on the encoding process persists over time, impairing long-term recall of negated information (Cornish and Wason, 1970; Craik and Tulving, 1975; Fiedler et al., 1996; Mayo et al., 2014; Zhang et al., submitted)[1]. For example, Mayo et al. (2014) reported the first comprehensive demonstration of the negation-induced amnesia effect. They found that actively negating a feature of an entity induced more memory loss of the entity itself compared to affirming the feature by conducting four tests in which they showed participants either short videos (Experiments 1–3) or verbal narratives (Experiment 4) embedded in a four-phase memory paradigm: study phase, verification task, distractive, and an incidental later free recall task. This negation-induced amnesia effect could be attributed to the short-term inhibitory effect of negation during the first memory test. Therefore, negation manipulates the encoding process to induce later forgetting in the retrieval phase.

Most research on the inhibitory effects of negation has been conducted with monolingual participants, while the processing of negation by bilinguals received little attention. Previous research on negation in bilinguals is generally driven by the idea that negation is universal and the processing of negation is more complicated than processing affirmation, regardless of the language (e.g., Hasegawa et al., 2002). Yet, since bilinguals have constant exercise to regulate the two languages they use, showing a stronger ability to resolve response conflict in non-linguistic activities (Bialystok et al., 2004; Costa et al., 2008), the study of negation processing in bilinguals may shed lights on the underlying mechanism of negation. To this end, the current study aimed to investigate the inhibitory effect of negation in a memory retrieval paradigm for bilinguals.

This study aims to explore the impact of language-switching on the processing of negation in an encoding-retrieval memory task. To this end, an online behavioral experiment was conducted with two groups of unbalanced Spanish-English bilinguals. One group of participants initially read stories in Spanish (L1), and the other group read the same stories in English (L2). Immediately after reading each story, the participants received a set of verification questions about the story contents, requiring a "yes" or "no" response. Some of the questions for verification were presented in the original language of the story (non-switch condition) and others in the alternative language (switch condition). In other words, the two critical manipulations of response polarity (affirmative vs. negative) and language sequence (switch vs. non-switch) occur in the verification tasks, given an opportunity to explore their combined effects on performance. Based on the literature reviewed above, we can expect both a switch cost and a negation cost in terms of longer response times and reduced accuracy. Most importantly, an interaction between the two factors is possible; for instance, the cost of negation could be reduced (primed) or increased (interfered) in the context of language switch, compared to the language non-switch. If so, this would suggest that the two processes share resources from the same inhibitory control mechanism.

Methods

Participants

A total of 121 psychology students from the University of La Laguna voluntarily participated in the current study. All the participants were neurologically healthy and right-handed with normal or corrected-to-normal eyesight. They were given informed consent and received course credit for their participation. Spanish is their native language (L1) and they use English as the second language (L2). Three participants were excluded for choosing "I find it difficult to understand most of the sentences." In a post-survey. The final sample consisted of 118 participants (98 females, $M = 20.4$ years, SD = 5.12).

To assess the participants' language proficiency, we inquired about the age of L2 acquisition (AoA), and administered a self-rated language skills questionnaire, in which participants rated on a five-point scale their own-perceived L1 (Spanish) and L2 (English) knowledge, with 5 indicating excellent and 1, poor. All

1 Zang, A., Beltrán, D., Wang, H., González, K. R., and de Vega, M. (submitted). *Does negation-induced forgetting result from inhibition or associative interference?* Available online at: http://dx.doi.org/10.2139/ssrn.4108574.

participants reported having an L2 level higher than the B1 in the CEFR test or an equivalent level in other English tests. As illustrated in **Table 1**, the self-rated questionnaire confirmed that the participants were unbalanced bilinguals with significantly higher proficiency in Spanish, than in English (t = 19.811, p < 0.001). The average age of L2 acquisition (AoA) was 5.05-year-old.

Materials

The experimental task was composed of a study phase and a verification phase. The study phase required participants to read four stories involving two protagonists (Montse and Jordi), describing their main personal traits and their activities in four different scenarios: daily life in the university, vacations, going to the beach, and a birthday party. Each story included 44–46 items each (M = 45.25), among which, 36 were about the protagonists. These experimental items consisted of 18 semantically related pairs, with each member of a pair assigned to one of the protagonists (e.g., Montse studies psychology, Jordi studies computer sciences). The remaining items were fillers (M = 9.25) to make the story natural and coherent (e.g., Then Montse and Jordi met in the library to study for a while). There were two versions of the stories written in Spanish (L1) and English (L2), respectively, although with identical content.

The verification phase was composed of 104 "wh" questions in total. Each story was followed by 26 questions, 18 of which referred to the experimental items shown in the preceding story (e.g., "Who studies psychology?"), and were followed by the name of one of the characters in a separate frame (e.g., Montse). The participants had to judge whether the name was a correct answer to the question, pressing the "yes" or the "no" response button. The remaining 8 questions referred to the filler items (e.g., Where did Montse and Jordi meet to study?). Of the experimental questions, 12 were non-switching questions asked in the same language as the initial story, and 6 were switching questions asked in the other language. The filler questions were always formulated in the same language as the story. All the questions were presented in pseudo-random order. For each story block, the first two questions were always fillers. The switching questions were always followed by 2–4 non-switching questions. Within each context language group, we created 8 counterbalanced lists resulting from 1) the facts attributed to the protagonists in the stories; 2) the facts asked in the verification questions 3) the response polarity.

Design and procedure

The experimental design was composed of Language Sequence (switch vs non-switch), and Response Polarity (affirmative vs negative), as within-subject factors, and Context Language (L1 vs L2) as a between-subject factor. Non-switch questions were in the same language as the context story and were preceded by a question in that language (L1 → *L1*, in L1 context, or L2 → *L2* in an L2 context), while the switch questions were in a different language from the context and were preceded by at least 2 questions in the context language (L1, L1 → *L2*, in the context of L1, or L2, L2 → *L1*, in the context of L2).

Due to the COVID-19 situation, the experiment was programmed and conducted online, using the Psytoolkit toolkit (Stoet, 2010, 2017). The participants were randomly and automatically assigned to the L1 or L2 story context. Fifty-three participants received most of the linguistic materials in Spanish (L1 Context), while the remaining 65 were assigned to English materials (L2 Context). A posterior test showed that the two context groups had similar language proficiency measures (see **Table 1**), according to the independent samples Mann–Whitney U tests: age of L2 acquisition [U(116) = 2,038.000, p = 0.083], L2 proficiency [Reading: U(116) = 2,058.500, p = 0.053, Writing: U(116) = 1,697.500, p = 0.887, Speaking: U(116) = 1,799.500, p = 0.660; Listening: U(116) = 1,941.500, p = 0.214; Average: U(116) = 1,929, p = 0.262] and L1 proficiency [Reading: U(116) = 1,618.500, p = 0.430; Writing: U(116) = 1,585.000, p = 0.389; Speaking: U(116) = 1,518.500, p = 0.157; Listening: U(116) = 1,620.500, p = 0.379; Average: U(16) = 1,527, p = 0.253].

Participants received an email with the experiment link and were instructed to complete the experiment online on a computer and a keyboard in a quiet room, previously turning off the mobile phone to avoid distractions. In the study phase, participants were first instructed to read the story carefully, keeping in mind that there would be related questions later. Then, the story was freely read by the participants in 4 self-paced paragraphs with 8–15 sentences in each paragraph (see **Supplementary material 3**). In the verification phase, each trial started with a 500 ms fixation in the center of the screen, followed by a question, which remained on the screen for 3,000 ms. Next, the protagonist's name was presented on the screen. Participants were prompted to press the "yes" response (the "j" key) or the "no" response (the "f" key) as fast and accurately as possible according to the initial story. If they failed to respond in the 5,000 ms, the program moved to the next sentence. The next trial started after a random blank period (1,000–1,200 ms). Participants were questioned on how well they had understood the story when they finished the experimental task. The questions were in a three-point scale: 1. I understand practically everything; 2. Moderate, I got lost with a few sentences; 3. Low, I find it difficult to understand most of the sentences.

Statistical analysis

To avoid language alternation influence on the non-switching level of the Language Sequence condition, the first one non-switching question following a switching question was

TABLE 1 Characteristics of participants.

	Group 1: Spanish context		Group 2: English context	
SELF-RATING	L1 (Spanish)	L2 (English)	L1 (Spanish)	L2 (English)
AOA		5.60 (2.88)		4.62 (1.74)
LISTENING	4.79 (0.45)	3.66 (0.95)	4.83 (0.48)	3.42 (1.02)
SPEAKING	4.58 (0.63)	3.26 (0.83)	4.71 (0.63)	3.18 (0.91)
READING	4.74 (0.48)	3.96 (0.80)	4.80 (0.44)	3.69 (0.78)
WRITING	4.47 (0.69)	3.13 (0.78)	4.60 (0.55)	3.17 (0.83)
MEAN	4.64 (0.58)	3.50 (0.90)	4.73 (0.54)	3.37 (0.92)

excluded from the statistical analysis. Nine participants and two items were excluded from the data analysis due to their high number of errors (> 40%). In addition, for each participant, verification trials with an incorrect response were excluded from the reaction times (RTs) analyses, as well as responses below 200 ms or above 2.5 standard deviations of the mean. Linear mixed-effect models (LMEMs) from the UllRtoolbox package were used to analyze the resulting trimmed RTs (R Core Team, 2015; Hernández, 2017), after normalizing with an inverse transformation (Box and Cox, 1964). Subjects and items were treated as random intercepts. Context Language, Language Sequence, and Response Polarity, as well as their interactions, were treated as fixed effects (Baayen et al., 2008). Before running the model, R-default treatment contrasts were automatically set to sum-to-zero contrasts. The structure of the estimated model employed to analyze the fixed-effects was: mod1.p = RT.p ∼ context language*response polarity*question language + (question language | sujeto) + (1 | item). More complex models including all relevant random structures were used in our initial analyses, but the models with more complex random structures failed to reliably converge (Barr, 2013). We called the Car package (Fox and Weisberg, 2018) with the function *car: Anova (χ2 variant)* to test significance and compute p values for the fixed-effects, avoiding issues of estimating denominator degrees of freedom in unbalanced designs, both mathematical and computational [see Alday et al. (2017), for an overview on parameter estimations and model fitting of LMEMs]. Since non-normality affects only the estimate of standard errors (and hence the significance of the contrasts), but not the fixed effects, a model using raw RTs was employed to extract mean differences to conduct *post hoc* contrasts.

For accuracy data, logistic regression models were estimated using as well the UllRtoolbox package (R Core Team, 2015; Hernández, 2017). Again, subjects and items were treated as random intercepts, while Context language, Language Sequence and Response Polarity, as well as their interactions, were treated as fixed effects (Baayen et al., 2008). The model used to analyze the fixed effects had the following structure: mod.hit = accuracy ∼ context language*response polarity*question language + (response polarity| sujeto) + (1| item). Type III Wald chi-square tests were adopted to test for significance and to calculate p values.

Results

The average RTs for correct trimmed response time and the percentage accuracy rates across conditions are shown in **Figure 1** and **Table 2**. The RTs analysis revealed significant effects of Response Polarity [$\chi2(1) = 381.50$, $p < 0.0001$], Language Sequence [$\chi2(1) = 8.30$, $p = 0.004$], and Context Language [$\chi2(1) = 5.43$, $p = 0.020$]. These effects reflected longer RTs for: 1) negation than affirmation responses, 2) switch than non-switch questions, and L1 than L2 contexts. The two-way interactions between Response Polarity and Language Sequence did not reach significance [$\chi2(1) = 1.55$, $p = 0.213$]. However, Context Language interacted significantly with Response Polarity [$\chi2(1) = 6.76$, $p = 0.009$] and with Language Sequence [$\chi2(1) = 23.75$, $p < 0.0001$]. *Post hoc* analyses revealed larger costs for language switch (the difference between switch and non-switch trials) when L1 was the main language (L1 context) ($\beta = -152.66$, SE = 28.4, $z = -5.370$, $p < 0.0001$) than when it was L2 ($\beta = -1.79$, SE = 26.7, $z = -0.067$, $p = 0.947$). Similarly, negation cost was larger in L1 context ($\beta = -193$, SE = 18.6, $z = -10.393$, $p < 0.0001$) than L2 context ($\beta = -108$, SE = 16.9, $z = -6.429$, $p < 0.0001$). More importantly, the three-way interaction between Response Polarity, Language Sequence and Context Language reached also significance ($\chi2(1) = 5.29$, p = 0.021) (See **Figure 1A**). Given the significant three-way interaction, our initial interest was to examine how Language Sequence and Response Polarity are processed in the Context Language L1 and L2. *Post hoc* analyses showed that, for the non-switch sequence, responding "yes" took similar time in the L1 Context and in the L2 Context; however, responding "no" took longer in L1 Context than in L2 Context (Affirmative: $\beta = -0.309$, SE = 1.39, $z = -0.222$, $p = 0.824$; Negative: $\beta = -3.146$, SE = 1.40, $z = -3.642$, $p = 0.0003$). Regarding the switch sequence, the two Context Languages differed significantly both when producing a "yes" and a "no" response (Affirmative: $\beta = -5.133$, SE = 1.44, $z = -3.571$, $p = 0.0004$; Negative: $\beta = -5.269$, SE = 1.45,

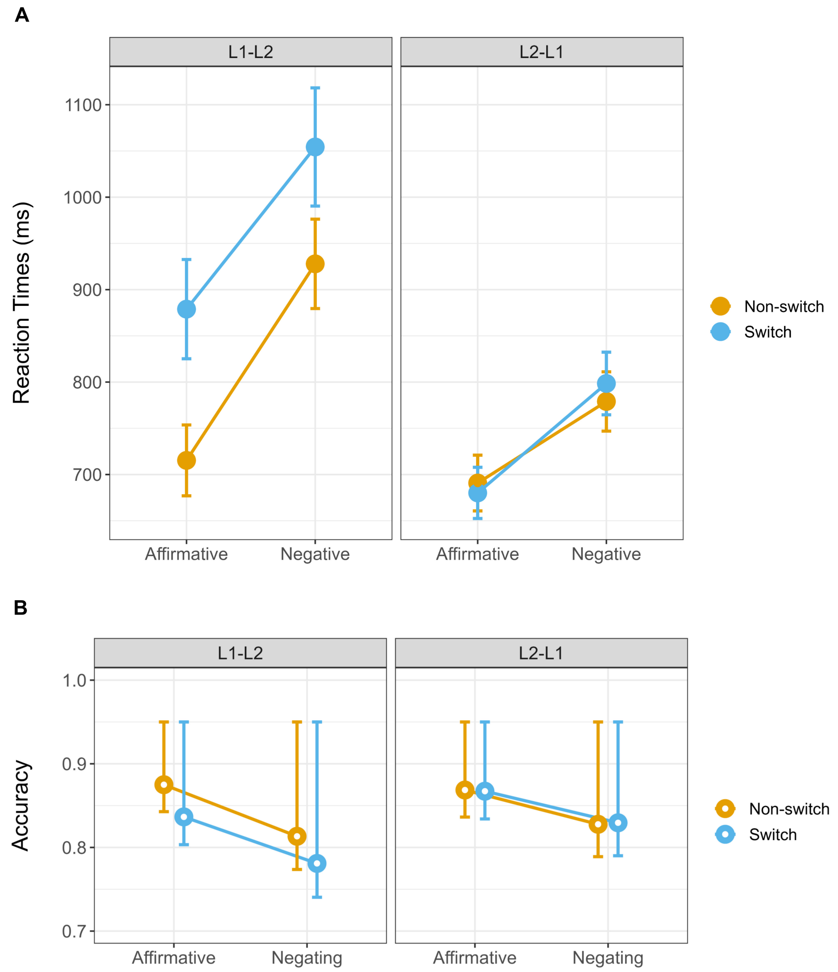

FIGURE 1
Mean RTs **(A)** and accuracy **(B)** for response polarity (affirmative vs. negative) language sequence (switch vs. non-switch) and context language (L1 vs. L2).

TABLE 2 Mean RTs (ms) and ACCs (%) in the Affirmative and Negative responses deposed by Language Sequence and Context Language.

		L1 Context Language		L2 Context Language	
		Switch	Non-switch	Switch	Non-switch
RT	Affirmative	883 (597)	716 (439)	677 (340)	694 (375)
	Negative	1,057 (681)	928 (539)	801 (405)	790 (391)
ACC	Affirmative	83.7 (37.0)	87.5 (33.1)	86.7 (34.0)	86.9 (33.8)
	Negative	78.1 (41.4)	81.3 (39.0)	82.9 (37.6)	82.8 (37.8)

$z = -3.642$, $p = 0.0003$), with a longer response time in L1 Context compared to L2 Context.

To better understand the three-way interaction, further analyses with the same LMEMs strategy as above was performed for each Context Language group separately. Fixed effects were reported for each Context Language group. For the L2 context group, there was a main effect of Response Polarity ($\beta = -4.741$, SE = 0.526, $t = -9.010$, $p < 0.0001$), indicating longer RTs for "no" than "yes" responses. However, for this group, neither Language Sequence ($\beta = 0.149$, SE = 0.731, $t = 0.205$, $p = 0.838$) nor the interaction reached significance ($\beta = -0.611$, SE = 0.750, $t = -0.815$, $p = 0.415$). In contrast, for the L1 context group, there were main effects of Response Polarity ($\beta = -7.638$, SE = 0.647, $t = -11.806$, $p < 0.0001$), and Language Sequence ($\beta = -4.802$, SE = 0.855, $t = -5.620$, $p < 0.0001$), which reflected longer RTs for negation than affirmation, and for switch than non-switch questions. Most importantly, the interaction between these two factors was significant for the L1 context group (β = 2.328, SE = 0.932, t = 2.499, p = 0.013), indicating relatively diminished switch cost for "no" responses in comparison to "yes" responses. All these findings indicated that the RT did not differ at the "baseline" condition (responding "yes" to non-switch questions), but the response time was highly and differentially modulated by Response Polarity and Language Sequence in both L1 and L2 groups.

For the accuracy rate logistic regression model, only the main effect of polarity resulted in significant, [$\chi 2(1) = 12.79$, $p < 0.01$], with higher accuracy for "yes" than "no" responses (see **Figure 1B**). The absence of the main effects of Language Sequence and Context Language as well as the interactions might be attributed to certain ceiling effect. The accuracy performance of encoding and retrieving the initial stories were similar for both groups of participants due to the ease of understanding the materials. Tables resuming the main results are viewable in **Supplementary Tables 1,2**.

Discussion

The present study investigated for the first time whether bilinguals' language switch modulates negation processing. To this end, we tested bilingual Spanish-English speakers using a two-step encoding and verification memory-based task. The encoding phase involved reading stories that were shown in the participants' native language (Spanish), or in their second language (English). The language switch was induced during the verification phase, by presenting the questions in the same language (non-switch) or in the alternative language (switch) as the main story. On the other hand, responses to both types of verification questions were affirmative (yes) or negative (no), so the polarity was not a feature of the sentences themselves, but rather arose during the response production. An unexpected result was that the verification times for the L1 Context were longer than for the L2 Context, which seems to be at odds with the well-known fact that unbalanced bilinguals are usually more efficient at processing their native language rather than a second language. However, this result is misleading if we neglect the interactive effects of Language Context with Response polarity and Language Sequence. Thus, if we focus on the baseline condition (non-switch and affirmative) the response time and accuracy did not differ for both language contexts. However, beyond the baseline, the RTs were modulated differently in the L1 Context (both by switching and negative polarity) and in the L2 Context (only by negative polarity). The main results of the modulation were as follows. First, an asymmetric language switch effect was found. That is, in the context of L1 (Spanish), switching to L2 in the verification involved more cognitive cost (slower and less accurate responses) than keeping the same language (L1-L1). However, in the context of L2 (English) switching to L1 did not imply an additional cognitive cost compared to keeping the same language (L2-L2). Secondly, a classical negation effect was observed, with a longer reaction time to produce negative ("no") than affirmative responses ("yes"). Finally, although the negation effect was patent for the L1 and L2 context groups, negation only interacted with language switching in the former. That is, in the context of L1, the cost of switching from L1 to L2 was reduced for "no" responses in comparison to "yes" responses, implying a sort of priming, which is, producing negative responses benefits from a language switching sequence. These results will be discussed in detail below.

Asymmetry of language switch cost

The finding of switching costs from L1 to L2 is consistent with the research on language-dependent differences in memory retrieval processes (Marian and Neisser, 2000; Marian and Fausey, 2006; Marian and Kaushanskaya, 2007; Wang, 2022). Such research revealed that when information was encoded in L1, and retrieval was in L2 (switch), the recall was impaired compared to using L1 in retrieval (non-switch). However, the switching cost was attenuated (found in accuracy but not speed) when encoding was in L2 and retrieval in L1 (switch) compared to using L2 both at encoding and retrieval (non-switch) (Larsen et al., 2002; Matsumoto and Stanny, 2006; Marian and Kaushanskaya, 2007). Another important finding in this literature is that proficiency moderates the effect on retrieval speed. Unbalanced bilinguals who are more proficient in one of the languages show a stronger switching effect when encoding is in the dominant language, with no language switch effect when encoding in the second language. Overall, this latency pattern coincided with that obtained in our study, with a sample of unbalanced bilinguals.

The switch cost asymmetry might be attributed to the usual direction of translation. As suggested by Marian and Kaushanskaya (2007), unbalanced bilinguals are more likely to mentally translate the less proficient language into the more proficient language (Dornic, 1978; Schrauf, 2002), resulting in stronger connections from L2 to L1 than from L1 to L2 (Kroll et al., 2002). In our case, switching from L2 to L1 was consistent with the most proficient direction of translation, resulting in the absence of language-dependent effects. A complementary explanation is that unbalanced Spanish-English participants, like ours, have more linguistic experience in retrieving information in Spanish from an English source than retrieving in English from a Spanish source, presenting an advantage for the verification in the L2-L1 direction in comparison to the L1-L2 direction (Marian and Fausey, 2006). In sum, our results confirmed a strong switching cost in the L1-L2 direction, and the absence of switching cost in the direction L2-L1 in a memory-based language-switching task.

The asymmetry of switching cost obtained here (poorer performance when switching from L1 to L2 than from L2 to L1) contrasts with the commonly reported pattern in language switching studies using naming tasks (Meuter and Allport, 1999). In these cases, the language switch asymmetry for unbalanced bilinguals was the opposite; that is, higher language-switching cost from L2 to L1 than the other way around. This is because naming in the non-dominant language requires more active inhibition on the dominant language and exerts negative priming on the following L1 trial. In contrast, little suppression is needed in the reversed direction (Green, 1998; Meuter and Allport, 1999; Finkbeiner et al., 2006; Declerck and Philipp, 2015a). Notably, the highest L1-L2 switching cost was mainly found in the lexical tasks of picture naming, while the highest L2-L1 switching cost, as the reported in the present study, is usually based on two-step memory tasks, involving more complex linguistic materials during the encoding phase (stories) and the delayed memory phase (sentence verification). In fact, numerous empirical studies have demonstrated that task-dependent factors might influence language transition cost, such as sentential context (Declerck and Philipp, 2015b), contextually changing language proficiency (Bonfieni et al., 2019), grammatical structure (Gollan and Goldrick, 2016), or cue-to-stimulus intervals (Verhoef et al., 2009), etc.

Language switch and negation

Consistent with the previous work (Clark and Chase, 1972; Carpenter and Just, 1975; Zhang et al., see text footnote 1), we found lower accuracy and longer RTs for negative than affirmative responses in both L1 and L2 context groups, suggesting that more elaboration and more cognitive resources were required for producing negative responses compared to producing affirmative responses. However, the major finding of the present study was the interaction between language sequence (switch vs. non-switch) and response polarity (affirmative vs. negative) in the most demanding switching context (from L1 to L2). Specifically, results showed that the switching cost from L1 to L2 diminished for negative responses compared to affirmative responses, indicating a priming of directional switch over negation. A statistical interaction between two variables in a reaction time task, in this case language switching and linguistic negation, may indicate that the processes underlying these variables share neurocognitive resources (Sternberg, 1998). In fact, there is independent evidence that language switch in bilinguals and processing negation utilize the general-domain mechanisms of inhibitory control.

Language switch has been described as a conflict monitoring process, since the bilinguals must be able to actively inhibit one language while using the other, to minimize interference. Neuroimaging studies have provided evidence that inhibitory control networks, including the anterior cingulate, the SMA or the prefrontal gyrus are recruited during language switch (Guo et al., 2011; Abutalebi et al., 2012; de Bruin et al., 2014). Moreover, in recent years, there is an emerging view that negation causes conceptual suppression by recruiting inhibitory mechanisms, particularly those concerned with preventing or stopping dominant reactions and representations (de Vega et al., 2016; Beltrán et al., 2019; Liu B. et al., 2020; Dudschig et al., 2021; Montalti et al., 2021). More relevant to our study, the denial function of negation is also empirically associated with inhibition effects (Cornish and Wason, 1970; Craik and Tulving, 1975; Fiedler et al., 1996; Mayo et al., 2014; Zhang et al., see text footnote 1). Specifically, the production of correct negative responses in the verification phase of a memory task impairs the long-term memory of the negated contents compared to the production of correct affirmative responses. The underlying mechanism of this amnesia effect of denial could also be attributed to the recruitment of inhibition, similar to the case of sentential negation (Mayo et al., 2014; Zhang et al., see text footnote 1). Based on these two accounts, we can interpret the reduced cost of negation in the context of language switching as supporting the idea that the two processes recruit the same neurocognitive mechanism of inhibitory control, producing a kind of priming effect. In other words, a question switching to the target (especially from L1 to L2) induces a strong inhibition state that could facilitate the inhibition-demanding negative responses. Hence, there is no need to reactivate the mechanism, and the negating response was facilitated. Note that we examined here the polarity effects in the production of affirmative or negative responses, rather than the comprehension of sentences differing in polarity, as frequently is done in other studies on the inhibitory effect of the negation (de Vega et al., 2016; Beltrán et al., 2018, 2019; Liu B. et al., 2020). It is possible that the language-switching priming effect on negation is confined to the production of negative responses, whereas no such effect would be obtained for the comprehension of negative sentences. The issue of how language switching and sentential negation influence each other requires further investigation. Although this measure did not provide unequivocal evidence that inhibition is the only mechanism under language control and negation, it shows the possibility that inhibition may explain at least part of the shared mechanism of language control processing and negation processing.

Conclusion and further avenues

This is the first study, to our knowledge, that examined how two apparently unrelated linguistic processes (language switching and producing negative answers) modulate each other in different language contexts. The choice of the two processes was motivated by the hypothesis that they recruit inhibitory control resources, and therefore they could interact when combined in the same task. The results found asymmetric switch cost (L1 to L2 > L2 to L1), negation cost (negation > affirmation), and interactive effects between them, which are suggestive of shared processes. This could have implications for theoretical and applied research fields, for instance, implement methods to learn a second language, better understand decision making processes, study inhibitory control disorders, long-term memory processes, etc. However, the current study has some limitations. First, despite the fact that the two groups of participants submitted to the L1 and L2 contexts, respectively, did not differ significantly in language proficiency and despite adopting the same material contents in Spanish and English for the study phase and verification phase, the between-group design of language context could induce biased results. Future studies are needed to adopt a within-participant design to better control for these possible biases. Second, future studies are needed to clarify whether the observed interactions between response polarity and language switch involves specific inhibitory control networks in the brain (e.g., SMA, rIFG, anterior cingulate cortex), using neuroimaging, EEG and non-invasive brain stimulation. Also, it might be useful to test these interactive effects with different task demands and materials, including sentential negation, naming paradigms, etc.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Comité de Ética de la Investigación y Bienestar Animal Vicerrectorado de Investigación y Transferencia de Conocimiento Universidad de La Laguna, La Laguna. Email: ceiba@ull.es. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AZ, MV, and DB contributed to the conception and design of the study. AZ and YF performed the statistical analysis. AZ wrote the first draft of the manuscript. MV and DB wrote sections of the manuscript. AZ, MV, YF, HW, and DB contributed to the manuscript revision, read, and approved the submitted version. All authors contributed to the article and approved the submitted version.

Funding

This work was co-funded by the Spanish MINECO and the European Regional Development Funds (Grant Number RTI2018-098730-B-100-R to DB and MV), and by the National Social Science Foundation of China (Grant Number 21FYYA002).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.906154/full#supplementary-material

References

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex* 22, 2076–2086. doi: 10.1093/cercor/bhr287

Abutalebi, J., and Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. Lang. Cogn. Process. 23, 557–582.

Alday, P. M., Schlesewsky, M., and Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: event- related potentials reflect continuous model updates. *ENeuro* 4:ENEURO.0311–16.2017, doi: 10.1523/ENEURO.0311-16.2017

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., et al. (2012). Grip force reveals the context sensitivity of language-induced motor activity during "Action Words" processing: evidence from sentential negation. *PLoS One* 7:e50287. doi: 10.1371/journal.pone.0050287

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Front. Psychol.* 4:328. doi: 10.3389/fpsyg.2013.00328

Bartoli, E., Tettamanti, A., Farronato, P., Caporizzo, A., Moro, A., Gatti, R., et al. (2013). The disembodiment effect of negation: negating action-related sentences attenuates their interference on congruent upper limb movements. *J. Neurophys.* 109, 1782–1792. doi: 10.1152/jn.00894.2012

Beltrán, D., Liu, B., and de Vega, M. (2021). Inhibitory mechanisms in the processing of negations: a neural reuse hypothesis. *J. Psychol. Res.* 50, 1243–1260. doi: 10.1007/s10936-021-09796-x

Beltrán, D., Morera, Y., García-Marco, E., and De Vega, M. (2019). Brain inhibitory mechanisms are involved in the processing of sentential negation, regardless of its content. Evidence from EEG theta and beta rhythms. *Front. Psychol.* 10:1782. doi: 10.3389/fpsyg.2019.01782

Beltrán, D., Muñetón-Ayala, M., and de Vega, M. (2018). Sentential negation modulates inhibition in a stop-signal task. Evidence from behavioral and ERP data. *Neuropsychologia* 112, 10–18. doi: 10.1016/j.neuropsychologia.2018.03.004

Bialystok, E., Klein, R., Craik, F. I. M., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Bloom, L. M. (1970). *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Bonfieni, M., Branigan, H. P., Pickering, M. J., and Sorace, A. (2019). Language experience modulates bilingual language control: the effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychol.* 193, 160–170. doi: 10.1016/j.actpsy.2018.11.004

Box, G. E. P., and Cox, D. R. (1964). An Analysis of Transformations. *J. R. Statist. Soc.* 26, 211–243. doi: 10.1111/j.2517-6161.1964.tb00553.x

Bultena, S., Dijkstra, T., and Van Hell, J. G. (2015). Language switch costs in sentence comprehension depend on language dominance: evidence from self-paced reading. *Bilingualism* 18, 453–469. doi: 10.1017/S1366728914000145

Carpenter, P. A., and Just, M. A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychol. Rev.* 82, 45–73. doi: 10.1037/h0076248

Choi, S. (1988). The semantic development of negation: a cross-linguistic longitudinal study. *J. Child Lang.* 15, 517–531. doi: 10.1017/S030500090001254X

Clark, H. H., and Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cogn. Psychol.* 3, 472–517. doi: 10.1016/0010-0285(72)90019-9

Cornish, E. R., and Wason, P. C. (1970). The Recall of Affirmative and Negative Sentences in an Incidental Learning Task. *Quart. J. Exp. Psychol.* 22, 109–114. doi: 10.1080/00335557043000032

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511. doi: 10.1016/j.jml.2004.02.002

Craik, F. I., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol.* 104, 268–294.

Davies, G. M., and Thomson, D. M. (1988). *Memory in Context: Context in Memory*. Hoboken: John Wiley & Sons.

De Bot, K. (1992). A bilingual production model: levelt's 'speaking'model adapted downloaded from. *Appl. Ling.* 13, 1–24.

de Bruin, A., Roelofs, A., Dijkstra, T., and FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: FMRI evidence from trilingual speakers. *NeuroImage* 90, 348–359. doi: 10.1016/j.neuroimage.2013.12.049

de Vega, M., Morera, Y., León, I., Beltrán, D., Casado, P., and Martín-Loeches, M. (2016). Sentential negation might share neurophysiological mechanisms with action inhibition. Evidence from frontal theta rhythm. *J. Neurosci.* 36, 6002–6010. doi: 10.1523/JNEUROSCI.3736-15.2016

Declerck, M., and Philipp, A. M. (2015b). A sentence to remember: instructed language switching in sentence production. *Cognition* 137, 166–173. doi: 10.1016/j.cognition.2015.01.006

Declerck, M., and Philipp, A. M. (2015a). A review of control processes and their locus in language switching. *Psychon. Bull. Rev.* 22, 1630–1645. doi: 10.3758/s13423-015-0836-1

Declerck, M., Philipp, A. M., and Koch, I. (2013). Bilingual control: sequential memory in language switching. *J. Exp. Psychol.* 39, 1793–1806. doi: 10.1037/a0033094

Dornic, S. (1978). "The bilingual's performance: Language dominance, stress, and individual differences," in *Language Interpretation and Communication*, eds D. Gerver and H. W. Sinaiko (New York: Plenum), 259–271.

Dudschig, C., Kaup, B., Svaldi, J., and Gulewitsch, M. D. (2021). Negation processing in children with ADHD: the generic problem of using negation in instructions. *J. Psychol. Res.* 50, 1309–1320. doi: 10.1007/s10936-021-09789-w

Fiedler, K., Walther, E., Armbruster, T., Fay, D., and Naumann, U. (1996). Do you really know what you have seen? Intrusion errors and presuppositions effects on constructive memory. *J. Exp. Soc. Psychol.* 32, 484–511. doi: 10.1006/jesp.1996.0022

Finkbeiner, M., Almeida, J., Janssen, N., and Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *J. Exp. Psychol.* 32, 1075–1089. doi: 10.1037/0278-7393.32.5.1075

Fox, J., and Weisberg, S. (2018). Visualizing fit and lack of fit in complex regression models: effect plots with partial residuals. *J. Statist. Softw.* 87, 1–27. doi: 10.18637/jss.v087.i09

Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., and Koch, I. (2021). Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – a meta-analysis. *J. Cogn.* 4:55. doi: 10.5334/JOC.186

García-Marco, E., Morera, Y., Beltrán, D., de Vega, M., Herrera, E., Sedeño, L., et al. (2019). Negation markers inhibit motor routines during typing of manual action verbs. *Cognition* 182, 286–293. doi: 10.1016/j.cognition.2018.10.020

Gollan, T. H., and Goldrick, M. (2016). Grammatical constraints on language switching: language control is not just executive control. *J. Mem. Lang.* 90, 177–199. doi: 10.1016/j.jml.2016.04.002

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism* 1, 67–81. doi: 10.1017/s1366728998000133

Guo, T., Liu, H., Misra, M., and Kroll, J. F. (2011). Local and global inhibition in bilingual word production: FMRI evidence from Chinese–English bilinguals. *NeuroImage* 56, 2300–2309. doi: 10.1016/j.neuroimage.2011.03.049

Hasegawa, M., Carpenter, P. A., and Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage* 15, 647–660. doi: 10.1006/nimg.2001.1001

Hernández, J. (2017). *ULLRToolbox for R (Version 1.0). [Software]*. Available online at: https://sites.google.com/site/ullrtoolbox/00-instalacion-windows (accessed November 10, 2020).

Kaup, B., and Zwaan, R. A. (2003). Effects of Negation and Situational Presence on the Accessibility of Text Information. *J. Exp. Psychol.* 29, 439–446. doi: 10.1037/0278-7393.29.3.439

Kroll, J. F., Michael, E., Tokowicz, N., and Dufour, R. (2002). The development of lexical fluency in a second language. *Sec. Lang. Res.* 18, 137–171. doi: 10.1191/0267658302sr201oa

Kroll, J. F., Dussias, P. E., Bice, K., and Perrotti, L. (2015). Bilingualism, mind, and brain. *Annu. Rev. Linguist.* 1:377.

Larsen, S. F., Schrauf, R. W., Fromholt, P., and Rubin, D. C. (2002). Inner speech and bilingual autobiographical memory: a Polish-Danish cross-cultural study. *Memory* 10, 45–54. doi: 10.1080/09658210143000218

Liu, B., Gu, B., Beltrán, D., Wang, H., and de Vega, M. (2020). Presetting an inhibitory state modifies the neural processing of negated action sentences. An ERP study. *Brain Cogn.* 143:105598. doi: 10.1016/j.bandc.2020.105598

Liu, H., Tong, J., de Bruin, A., Li, W., He, Y., and Li, B. (2020). Is inhibition involved in voluntary language switching? Evidence from transcranial direct current stimulation over the right dorsolateral prefrontal cortex. *Int. J. Psychophysiol.* 147, 184–192. doi: 10.1016/j.ijpsycho.2019.12.002

MacDonald, M. C., and Just, M. A. (1989). Changes in Activation Levels With Negation. *J. Exp. Psychol.* 15, 633–642. doi: 10.1037/0278-7393.15.4.633

Marian, V., and Fausey, C. M. (2006). Language-dependent memory in bilingual learning. *Appl. Cogn. Psychol.* 20, 1025–1047. doi: 10.1002/acp.1242

Marian, V., and Kaushanskaya, M. (2007). Language context guides memory content. *Psychon. Bull. Rev.* 14, 925–933. doi: 10.3758/BF03194123

Marian, V., and Neisser, U. (2000). Language-dependent recall of autobiographical memories. *J. Exp. Psychol.* 129, 361–368. doi: 10.1037/0096-3445.129.3.361

Matsumoto, A., and Stanny, C. J. (2006). Language-dependent access to autobiographical memory in Japanese-English bilinguals and US monolinguals. *Memory* 14, 378–390. doi: 10.1080/09658210500365763

Mayo, R., Schul, Y., and Burnstein, E. (2004). "I am not guilty" vs "I am innocent": successful negation may depend on the schema used for its encoding. *J. Exp. Soc. Psychol.* 40, 433–449. doi: 10.1016/j.jesp.2003.07.008

Mayo, R., Schul, Y., and Rosenthal, M. (2014). If you negate, you may forget: negated repetitions impair memory compared with affirmative repetitions. *J. Exp. Psychol.* 143, 1541–1552. doi: 10.1037/a0036122

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40. doi: 10.1006/JMLA.1998.2602

Montalti, M., Calbi, M., Cuccio, V., Umiltà, M. A., and Gallese, V. (2021). Is motor inhibition involved in the processing of sentential negation? An assessment via the Stop-Signal Task. *Psychol. Res.* 0123456789. [Epub ahead of print] doi: 10.1007/s00426-021-01512-7

Orenes, I., Beltrán, D., and Santamaría, C. (2014). How negation is understood: evidence from the visual world paradigm. *J. Mem. Lang.* 74, 36–45. doi: 10.1016/j.jml.2014.04.001

Papeo, L., Hochmann, J. R., and Battelli, L. (2016). The default computation of negated meanings. *J. Cogn. Neurosci.* 28, 1980–1986. doi: 10.1162/jocn_a_01016

Pea, R. D. (1980). "The development of negation in early child language," in *The Social Foundations of Language and thought:Essays in Honor of Jerome S. Bruner*, ed. D. R. Olson (New York, NY: W. W. Norton), 156–186.

R Core Team (2015). *R: A language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schrauf, R. W. (2002). Bilingual inner speech as the medium of cross-modular retrieval in autobiographical memory. *Behav. Brain Sci.* 25, 698–699.

Sternberg, S. (1998). "Discovering mental processing stages: The method of additive factors," in *Methods, Models, and Conceptual Issues: An Invitation to Cognitive Science*, eds D. Scarborough and S. Sternberg (Cambridge: The MIT Press), 703–863.

Stoet, G. (2010). PsyToolkit: a software package for programming psychological experiments using Linux. *Behav. Res. Methods* 42, 1096–1104. doi: 10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: a novel web-based method for running online questionnaires and reaction-time experiments. *Teach. Psychol.* 44, 24–31. doi: 10.1177/0098628316677643

Terao, Y., and Ugawa, Y. (2002). Basic mechanisms of TMS. *J. Clin. Neurophysiol.* 19, 322–343. doi: 10.1097/00004691-200208000-0000

Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., et al. (2008). Negation in the brain: modulating action representations. *NeuroImage* 43, 358–367. doi: 10.1016/j.neuroimage.2008.08.004

Tomasino, B., Weiss, P. H., and Fink, G. R. (2010). To move or not to move: imperatives modulate action-related verb processing in the motor system. *Neuroscience* 169, 246–258. doi: 10.1016/j.neuroscience.2010.04.039

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373. doi: 10.1037/h0020071

Verhoef, K., Roelofs, A., and Chwilla, D. J. (2009). Role of inhibition in language switching: evidence from event-related brain potentials in overt picture naming. *Cognition* 110, 84–99. doi: 10.1016/j.cognition.2008.10.013

Vitale, F., Monti, I., Padrón, I., Avenanti, A., and de Vega, M. (2021). The neural inhibition network is causally involved in the disembodiment effect of linguistic negation. *Cortex* 147, 72–82. doi: 10.1016/j.cortex.2021.11.015

Wang, L. (2022). Influences of Language Shift on Speech Fluency in Memory Production of Unbalanced Chinese-English Bilinguals. *Theory Practice Lang. Stud.* 12, 375–381. doi: 10.17507/tpls.1202.21

Wang, X. (2015). Language control in bilingual language comprehension: evidence from the maze task. *Front. Psychol.* 6:1179. doi: 10.3389/fpsyg.2015.01179

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 January 2023
DOI 10.3389/fpsyg.2022.923123

OPEN ACCESS

EDITED BY
Huili Wang,
Dalian University of Technology, China

REVIEWED BY
Max Ryan Freeman,
St. John's University, United States
Eve Higby,
California State University, East Bay,
United States

*CORRESPONDENCE
Erik D. Thiessen
Thiessen@andrew.cmu.edu

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 18 April 2022
ACCEPTED 16 November 2022
PUBLISHED 05 January 2023

CITATION
Kim Y, Ye Z, Leventhal Z, Wang W-J and Thiessen ED (2023) Effects of language background on executive function: Transfer across task and modality. *Front. Psychol.* 13:923123.
doi: 10.3389/fpsyg.2022.923123

Effects of language background on executive function: Transfer across task and modality

Yeonwoo Kim[1], Zixuan Ye[2], Zachary Leventhal[2,3], Wei-Ju Wang[2] and Erik D. Thiessen[1]*

[1]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States, [2]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA, United States, [3]Department of Economics, Carnegie Mellon University, Pittsburgh, PA, United States

The relation between linguistic experience and cognitive function has been of great interest, but recent investigations of this question have produced widely disparate results, ranging from proposals for a "bilingual advantage," to a "bilingual disadvantage," to claims of no difference at all as a function of language. There are many possible sources for this lack of consensus, including the heterogeneity of bilingual populations, and the choice of different tasks and implementations across labs. We propose that another reason for this inconsistency is the task demands of transferring from linguistic experience to laboratory tasks can differ greatly as the task is modified. In this study, we show that task modality (visual, audio, and orthographic) can yield different patterns of performance between monolingual and multilingual participants. The very same task can show similarities or differences in performance, as a function of modality. In turn, this may be explained by the distance of transfer – how close (or far) the laboratory task is to the day to day lived experience of language usage. We suggest that embodiment may provide a useful framework for thinking about task transfer by helping to define the processes of linguistic production and comprehension in ways that are easily connected to task manipulations.

KEYWORDS

cognition, bilingual, executive function, switch costs, modality

Introduction

Perhaps due to the close interrelation between language and thought (e.g., Chomsky, 1964), researchers have long suspected a link between linguistic experience and cognitive abilities. However, the nature of this hypothesized link has varied widely across time and contexts. Early in the 20th century, researchers and educators in the United States confidently pronounced that monolingualism was the "correct" way to raise children, and any deviations from this standard increased the risk of developmental delays or disorders (e.g., Saer, 1923; Yoshioka, 1929). Within decades, the opinion among researchers changed dramatically, due to pioneering research by Peal and Lambert (1962), a tradition subsequently much expanded by Bialystok (2017) and

Bialystok et al. (2007) which provided evidence that bilingualism actually promotes beneficial cognitive outcomes in executive function skills, metalinguistic awareness, and cognitive flexibility, as well as evidence that bilingualism protects against cognitive decline and dementia (for review, see Adesope et al., 2010). In particular, the *cognitive reserve* framework has provided a conceptual account for understanding the protective effects of bilingualism, especially in aging (see Bialystok, 2021). This framework suggests that the protective effects of bilingualism result from the lifelong experience of resolving competition between the two (or more) jointly activated languages with which bilinguals are familiar. This results in a series of adaptations to cognitive and neurological systems creating conditions in which cognitive actions are more automatic and less effortful (Cabeza et al., 2018).

While much research has supported the cognitive reserve framework, or related bilingual cognitive advantages, there have been notable exceptions. Objections to the perspective have been made from both theoretical and empirical grounds. Bialystok (2017) have argued that the lifelong experience of selection between languages provides multilingual speakers with some generalized advantage in attention, inhibition, or selection. From this perspective, language selection is explicitly embedded within a domain-general process of executive function (e.g., Blumenfeld and Marian, 2013; Paap et al., 2018). However, the idea of a unified set of executive function processes is itself a complex and disputed claim. For example, there are few (if any) process-pure measures of executive function abilities; instead, tasks are likely to involve multiple executive function skills in different degrees (for discussion, see Miyake et al., 2000). This task impurity problem makes it difficult to generate *a priori* predictions about which tasks should, or should not, show effects of multilingual experience (for discussion, see Hartsuiker, 2015).

A related objection to the cognitive reserve framework is that the phenomenon itself is not robust or replicable. A number of investigators have reported failures to replicate the phenomenon of a multilingual advantage (e.g., Paap and Greenberg, 2013; Gathercole et al., 2014). These non-replications raise the possibility that any bilingual advantage is too small or unreliable to be of practical significance, or perhaps even an illusion. It should be noted that many of these non-replications of a bilingual advantage (though not all; e.g., Antón et al., 2016) have been obtained with undergraduate or healthy adult participants. The cognitive reserve framework explicitly argues that the "protective" or "beneficial" effects of bilingualism should be most apparent in situations where cognitive resources are heavily taxed or drained. Therefore, the framework predicts that healthy young adults are the group least likely to show the beneficial or protective effects of bilingual experience, which are thought to be more pronounced in children and older adults (e.g., Craik et al., 2010; Bialystok, 2021). Consistent with this, a recent meta-analysis (Ware et al., 2020) indicated that evidence for a bilingual advantage is stronger for participants over 50 than for participants between the ages of 18 and 29.

At minimum, then, it is apparent the bilingual advantage is not always easily detected, or is more easily observed in some contexts than others. As such, making *a priori* predictions about when one should (or should not) expect to find differences as a function of language background is of premium importance in theory testing. We propose that an important factor to consider in making these predictions is the "distance of transfer." Transfer refers to the process of executing some learned behavior or process in a novel context; for example, after learning to golf on one course, playing a new course requires transferring those learned golf skills to a new setting (for discussion, see Salomon and Perkins, 1989). In some cases, transfer of learning only requires adapting to superficial, perceptual differences in context or setting, which is referred to as "near" transfer. In other cases, learners might be asked to apply their skills in settings that are quite distinct from their training. For example, after being trained on a working memory task with numerical digits, being asked to perform a working memory task with orthographic numbers would be a case of near transfer, while being asked to remember chess displays would be a case of far transfer. Training in executive function skills—of the type invoked by the cognitive reserve framework and its critics—typically results in near transfer, rather than far transfer (Kassai et al., 2019; Sala and Gobet, 2019; Gobet and Sala, 2020). Evidence for far transfer is rare, a finding that appears to hold across the lifespan (Sala et al., 2019).

Consideration of transfer suggests that if there is an effect of bilingualism on executive function, we are much more likely to observe it in tasks or settings that are very similar to (multi-) language use, and less likely to observe it in contexts that are less similar to linguistic stimuli, tasks, or processes. Of course, this requires us to define some framework or property structure along which to evaluate similarity; in isolation, the notion of similarity is notoriously susceptible to circularity (Goodman, 1972; Tversky, 1977). To ground our perspective on similarity, we will rely on the embodiment theory of cognition, which argues that the mind's experience of cognition is deeply rooted in the body's interactions with the world (e.g., Barsalou, 1999a). That is, when comprehending or producing a verb like "open," humans do not rely on an abstract, symbolic, or propositional definition of the verb; instead, they recall their physical experiences with opening doors, drawers, or containers (Barsalou, 1999b). In this way, cognition is always contextually situated, and the critical factors that define similarity are the (embodied) representations invoked in a task, and the operations over those representations. This is in some ways antithetical to the claims of the domain-general executive function "reserve" invoked by the cognitive reserve framework (Bialystok, 2017). However, we believe it may be a better fit to the (somewhat contradictory) state of the literature, in part by providing an opportunity to generate explanations and predictions about the replications and non-replications of bilingual advantage in executive function.

From the embodiment perspective, the inconsistent pattern in the literature—with both replications and non-replications of a bilingual advantage—may be related not only to differences in the participant population (young adults, as opposed to children or older adults), but also due to differences in the distance of transfer between the experience of bilingualism and the laboratory task(s) used by specific laboratories or investigators. The current investigation was conducted to investigate the plausibility of this claim. In particular, we examined the prediction that the more "language-like" a task is, the more likely it is to show evidence of a difference in performance between monolingual and multilingual speakers. That is, the more language-like a task is, the more likely it should be to provide replicable evidence of some effect of linguistic background.

Our goal in this study is to assess whether evidence of effects of linguistic background on executive processing are more or less observable as a function of the linguistic nature of the task and stimuli. Because executive function is a broad umbrella concept, we particularly focus our attention on the construct of "response selection," which has often been described as a key shared component between multilingual experience and executive function tasks. The cognitive reserve framework argues that it is the lifelong experience of regularly being asked to disengage attention from the non-target information (i.e., language), and to switch attention to relevant information, that is the mechanism responsible for subsequent advantages in executive function (Bialystok, 2009). This is specifically an appeal to domain-general attentional processes, rather than to domain-specific (e.g., perceptual) inhibitory response (c.f. Freeman et al., 2017; Paap et al., 2018). As such, we will present participants with two response selection tasks that have been suggested to be related to the attentional control skills accentuated by multilingual experience: the Simon task (Simon and Small, 1969) and the dimension-switching task (Prior and MacWhinney, 2010).

In addition to assessing participants ability to perform conflict resolution in a relatively decontextualized laboratory task, we will also ask participants to perform conflict resolution in a (somewhat) more linguistically-relevant task: word segmentation (e.g., Saffran et al., 1996). The ability to segment words from fluent speech is a fundamental linguistic skill, and one that reflects a listener's fluency with the language. Because words in fluent speech are not consistently marked by pauses or any other acoustic feature (e.g., Cole and Jakimik, 1980), listeners must rely on probabilistic cues—such as phonological structure—to identify where words begin and end in fluent speech (e.g., Johnson and Jusczyk, 2001). Fluent speakers quickly and automatically integrate information across multiple probabilistic cues to identify the most likely word boundaries in an utterance (e.g., Nazzi et al., 2014). In some cases, because cues are probabilistic, they may indicate different word boundaries. For example, most content words in English are stressed on their first syllable, and as such English speakers treat lexical stress as a cue to word onset (e.g., Jusczyk et al., 1993). Nevertheless, in a word like *giraffe*, where stress falls on the second syllable, listeners must be able to rely on other cues to correctly segment the word.

As this example shows, probabilistic cues can conflict and compete with each other, and successful word segmentation involves resolving this competition. Prior research has found that the ability to perform conflict resolution in laboratory tasks is predictive of the ability to perform conflict resolution in the word segmentation task, and successfully learn words (Weiss et al., 2010). This relationship has been observed for both monolingual and multilingual speakers (Bartolotti et al., 2011). We will attempt to replicate this relationship in the current investigation, in particular following the Weiss et al. (2010) methodology. The fact that this replication involves a task—word segmentation—that is fundamentally grounded in real language comprehension and processing should make this an instance of relatively near transfer, and thus perhaps more replicable than other tasks.

Finally, as a direct test of the hypothesis that transfer distance matters, we will explicitly manipulate the linguistic nature of the stimuli used in the dimension-switching task. The majority of research investigating the effects of linguistic experience on conflict resolution has relied upon tasks that involve visual stimuli. Relatively less work has examined performance with auditory tasks, though some systematic investigations of stimulus and task modality have been attempted (e.g., Calabria et al., 2012; Foy and Mann, 2014; MacNamara and Conway, 2014). We hypothesize that language background will more strongly influence participants' performance in a task with linguistic stimuli (words) than in a task with non-linguistic stimuli (images). This hypothesis is based on prior evidence that training of executive function abilities largely results in close transfer (that is, improvements on tasks that are very similar to the training experience), and only more rarely gives rise to far transfer (e.g., Kassai et al., 2019). If experience with multiple languages does indeed train executive function, that training should have greater efficacy for tasks that are similar to the training—that is, tasks that involve linguistic stimuli or linguistic processes. Tasks that involve non-linguistic stimuli or processes should show less effect of prior linguistic background. To test this hypothesis, we will present participants with a dimension-switching task, one that has previously been used to investigate differences between monolingual and multilingual participants (e.g., Prior and MacWhinney, 2010). In this task, participants must switch between rating stimuli on one dimension (living/non-living) to rating stimuli on another dimension (large/small). Consistent with prior work, we expect that monolingual participants will show a larger switch cost (i.e., slower responses) when prompted to switch from one rating dimension to the other, while multilingual participants should show a smaller switch cost. Of particular interest to us is whether the advantage in switch cost for multilingual participants (compared to monolingual participants) differs as a function of whether the rated stimuli are words or images.

To investigate these questions, we recruited a large sample of undergraduate participants, and asked them to provide us with extensive information about their previous language usage. With this information, we sorted participants into categories (monolingual, bilingual, trilingual, etc.), as well as analyzed individual differences in language background as a continuous variable. All participants completed a battery of tests assessing working memory, for which we do not expect to find differences as a function of language background (e.g., Lehtonen et al., 2018; Antón et al., 2019); this battery provides some measure of information about whether our participant groups are well matched on dimensions other than language background. All of these participants then completed the Simon task, and a word segmentation task with either converging or conflicting cues (e.g., Weiss et al., 2010). Finally, participants completed the task-switching task with either word or image stimuli. Taken together, these tasks will provide us with insight on the replicability and domain generality of effects of linguistic background on executive function skills.

Materials and methods

Participants

Two hundred and eighteen introductory psychology course students enrolled at Carnegie Mellon University, aged 18–25 years, participated for class credit (141 female, 70 male, 5 nonbinary, 2 declined to disclose). We excluded 10 participants who did not complete all the tasks from the main analyses, resulting in 208 participants in our final sample. Of these 208 participants, 108 were of East Asian origin, 64 were White, 22 were South Asian, 18 were of Hispanic, Latino, or Spanish origin, 10 were Black or African American, 8 were Southeast Asian, and 2 were Middle Eastern or North African (Note: participants were allowed to select multiple races, resulting in a sum greater than 208). Fifty participants reported that they were monolingual English speakers. Of the remaining multilingual participants, 95 identified as bilingual, 50 identified as trilingual, and 13 participants rated themselves as familiar with four or more languages. Descriptive statistics on the age of acquisition, usage, and proficiency of each language reported are summarized in Table 1.

Because multilingual participants are necessarily heterogeneous, we supplemented our categorization of participants as "monolingual" or "multilingual" with a more continuous measure. In recent years, both proponents and skeptics of a bilingual advantage have made attempts to treat language experience as more of a continuum, and less of a categorical variable (e.g., Anderson et al., 2018; Paap et al., 2018; Anderson et al., 2020b). This is consistent with the arguments of embodied cognition, which suggest that abstract categorical variables like "bilingual" are less informative than detailed information about the contexts in which a person uses language, and the tasks they perform with that language. Fortuitously, this continuous perspective on language use is also often advantageous in terms of statistical power (e.g., Peña et al., 2016). By adopting a continuous perspective, we do not mean merely looking at the age at which acquisition of L1 and L2 occurred, or the number of years a participant has spent using their language(s). Rather, our embodiment perspective suggests we should focus on the tasks that language is used with and for. This is in many ways consistent with the "adaptive control hypothesis" (Green and Abutalebi, 2013), which predicts that the neural underpinnings of language control and processing should adapt to the control demands presented by the interaction between multiple known languages (for review, see Abutalebi and Green, 2016). In particular, we will focus on participants' experience reading (and writing), comprehending, and speaking languages. Information about participants' language use in these three task domains, as well as the context in which they use language, and the age and amount of use for each language, was collected via a previously validated survey, the Language Experiences and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007).

Stimuli

Word segmentation task

This task was designed following Weiss et al. (2010), but with some alterations to the artificial language on account of constraints of our speech synthesis device. Participants listened

TABLE 1 Descriptive statistics of language background (age of acquisition, usage, proficiency) for each language reported.

	LDom1 (*N* = 173)		LDom2 (*N* = 78)		LDom3 (*N* = 24)		LDom4 (*N* = 5)	
Age acquired	1.53	(1.69)	4.27	(3.84)	10.71	(3.97)	11.40	(5.08)
Age Fluent	5.06	(2.43)	9.91	(5.78)	14.04	(4.30)	14.00	(4.30)
Age Reading	4.39	(1.55)	7.50	(4.24)	11.71	(3.45)	11.40	(5.08)
Age Fluent Reading	6.79	(2.32)	11.29	(4.94)	13.46	(4.42)	14.20	(3.35)
Time in Country	17.87	(3.27)	4.96	(6.10)	2.53	(6.36)	4.00	(7.87)
Time with Family	17.00	(5.47)	8.34	(9.22)	1.66	(5.50)	4.40	(7.80)
Time in School Work	15.43	(3.82)	4.62	(4.21)	3.39	(5.44)	4.80	(7.82)
Reading Proficiency	9.63	(0.81)	7.15	(2.35)	6.29	(2.65)	5.40	(2.30)
Speaking Proficiency	9.76	(0.55)	7.08	(1.93)	5.42	(2.26)	4.00	(1.58)
Understanding Proficiency	9.73	(0.63)	8.18	(1.63)	6.33	(2.30)	5.00	(1.73)
TF score	12.97	(0.25)	10.94	(2.43)	8.92	(2.62)	5.80	(2.78)

These data are reported for those participants from whom we have complete data. Means are reported first, with standard deviations inside parentheses. Proficiency measures range from 0 to 10. TF score is the number of "True" responses for the 13 true–false questions assessing language familiarity.

TABLE 2 (A) Number of monolingual participants in each word segmentation task condition. (B) Number of multilingual participants in each word segmentation task condition.

Language condition	Pause duration (milliseconds) 0 ms	10 ms	15 ms	25 ms	50 ms	Total
(A) Monolingual						
Convergent	6	4	2	2	10	24
Conflicting	8	8	3	4	3	26
Total	14	12	5	6	13	50
(B) Multilingual						
Convergent	15	7	14	11	29	76
Conflicting	17	15	17	17	16	82
Total	32	22	31	28	45	158

TABLE 3 List of 12 objects used in task-switching task, categorized into Living/Nonliving and Small/Large.

	Living	Nonliving
Small	Ant, Lemon, Rose	Dice, Fork, Pencil
Large	Tree, Dolphin, Cow	Chair, House, Bed

to an artificial language composed of four bisyllabic words, bugo (/bu.goʊ/), dapu (/dæ.pu/), diti (/di.ti/) and dobi (/doʊ.bi/). To generate this language, consonant-vowel syllables were synthesized in isolation, at a monotone 230 Hz. These syllables were then concatenated into a sequence with no pauses between syllables, ordered such that each bisyllabic word occurred 90 times, and never followed itself. This artificial language has no acoustic cues to word boundaries, but participants can discover words by attending to the likelihood of syllable co-occurrence. Syllables within a word always predict each other; so, for example, when a participant hears "bu," they will always hear "go" next. At the end of a word, any of the other three words can occur, so the co-occurrence probabilities are much lower at word boundaries (Aslin et al., 1998).

After creating this artificial language, we modified it further to create two versions: a Convergent Cue version, and a Conflicting Cue version. In the Converging Cue version, pauses were added between the words. Thus, a participant might hear "diti (pause) bugo (pause) dobi (pause)....," with a pause marking the boundary between words, consistent with the statistical information about syllable co-occurrence.

In the Conflicting Cue version, pauses occurred in the middle of each bisyllabic word. Thus, participants might hear "di (pause) tibu (pause) godi (pause)...," with the pauses interrupting each (statistically defined) word. In this language, the cues to word boundaries conflict. Co-occurrence cues indicate one set of boundaries, while pause cues indicate a different set of boundaries.

Participants were assigned one of five pause duration conditions: 0 ms (at which duration the Converging Cue and Conflicting Cue languages are identical), 10 ms, 15 ms, 25 ms, and 50 ms (at which duration pilot testing indicated that the co-occurrence cue and the pause cue are of roughly equivalent strength). See Table 2 for the number of participants in each condition, divided into monolingual and multilingual participants.

The concatenated version of the language was approximately two minutes long, varying slightly as a function of pause duration. After listening to the language for two minutes, participants were presented with a set of test trials. On each test trial, participants heard both a word (either diti or dapu) from the language, and a syllable combination formed across word boundaries (godi, or tibu), which we call a part-word. Test items were presented with no pauses between syllables.

Dimension-switching task

In this task, participants were presented with either a series of images, or a series of words. Each item in the series was surrounded by either a red or a blue border. The border indicated whether participants should rate the presented item as living/nonliving, or as large/small (the mapping between color and rating task was counterbalanced across participants).

Image stimuli were photographs with backgrounds removed adapted from prior published work (Moreno-Martínez and Montoro, 2012). Objects from this collection of images were balanced between categories by number of letters and syllables. The stimuli were presented within a bounding box of 720 pixels by 540 pixels. The average width of an image was 362 pixels (min: 35 px, max: 674 px, SD: 177 px) and a height of 334 pixels (min: 81 px, max: 684 px, SD: 148 px). Word stimuli labeled the same set of concepts as presented in the image stimuli, and were displayed with the first letter capitalized in black, Open Sans font, on a white background.

Table 3 lists the 12 images and words that were depicted. Red and blue borders were added to both image and word stimuli, creating a total of 24 stimuli (each of the 12 items with either a red or a blue border).

Procedures

All participants completed the experiment remotely, through personally owned computers at times and places of their convenience. Access was restricted to computers (no phones or tablets were permitted), and the experiment was conducted via Gorilla, a web platform for experiments. All participants completed the experiment in English. Consent was collected and a sound check was conducted before any data were collected. Participants then completed the word segmentation task, followed by the Simon task, the audio and visual digit span tasks, and the N-back tasks in randomized order. Afterwards, participants completed questionnaires on language experience and demographics, as well as the dimension-switching task. All task instructions were presented visually on screen before the task began; participants were asked to press the spacebar to indicate that they understood the instructions and were ready to proceed.

Word segmentation task

Participants were randomly assigned to a cue version of the language (Convergent, Conflicting) and a pause duration (0 ms, 10 ms, 15 ms, 25 ms, 50 ms). In the Convergent Cue versions of the language, pauses occurred at statistically-defined word boundaries. In the Conflicting Cue versions of the language, pauses occurred between syllables within statistically defined-words (and thus, indicate different segmentation points than the statistical information). Note that at 0 ms, the Conflicting and Convergent Cue versions are identical, as the pauses do not occur (ie, they have a length of 0 ms).

After participants were randomly assigned to cue condition and pause duration, they listened to the appropriate version of the artificial language for approximately two minutes. After listening to this artificial language, participants were presented with eight test trials, in which they were asked to identify which item sounded more familiar to them. On each trial, participants heard two test items (one word, and one part-word), and were instructed to press the "1" key if they believed that the first item sounded more familiar to them, and the "2" key if they believed the second word was more familiar. Test item presentation order within and across trials was counterbalanced across participants.

Simon task

On each trial, the word "LEFT" or "RIGHT" was presented on either the left or right side of a fixation cross. Participants were instructed to click "Q" when the word "LEFT" appeared on the screen and "P" when the word "RIGHT" appeared, regardless of their location on the screen. The task consisted of 32 randomized trials with feedback. In each trial, the word "LEFT" or "RIGHT" was displayed for 900 ms, followed by a 500 ms pause before the next trial. Participants could respond at any time during the total 1,400 ms. Upon a correct response, the screen displayed "Correct," covering the fixation cross. If the participant responded incorrectly or did not respond within the 1,400 ms, the screen displayed "Incorrect," as well as a brief reiteration of the instructions. Feedback for both correct and incorrect trials were shown for 2,000 ms, either immediately after the response, or after the 1,400 ms maximum trial time, if the participant did not respond.

Audio and visual digit span tasks

A series of randomly generated numerical digits were presented one at a time. In the visual span, the digits were presented on the screen with a fixation cross in between each digit. In the audio digit span, the digits were read with a 1-s pause between each digit. Participants were instructed to focus on the fixation cross during the reading of the digits. The task was to memorize the digits in order and type them after they were presented. The trials began with a four-digit-trial ("1 2 3 4"), followed by 2 trials of each number of digits, from 4 to 9 digits, for a total of 13 trials.

N-back task (2-back)

A series of letters were displayed on the screen one at a time in a pseudo-randomized order. Participants were instructed to indicate whether the current letter was the same as the one that appeared two letters ago, by pressing either "F" or "J" on the keyboard. The key that indicated a match was counterbalanced between participants. The task was divided into three blocks of 10 trials each, with the instructions redisplayed on the screen between each block.

Language and demographics questionnaires

Participants completed the LEAP-Q (Marian et al., 2007) regarding their linguistic background and experience. Familiarity with a language was determined using 7 free-response questions regarding age and setting of acquisition. Proficiency and extent of use for each language with which the participant reported familiarity was assessed using three likert scales, and 13 true-false questions, for each language (up to 4). After completing questions about their linguistic background, participants were asked to provide general demographic information such as age and gender, as well as ethnicity and nationality.

Dimension-switching task

In the dimension-switching task, participants were presented with a stimulus (either an image or a word), surrounded by a solid colored border. The color of this border indicated the judgment the participants should make about the stimulus: whether it is living/nonliving, or larger/smaller than a shoebox (Figure 1). Each participant was randomly assigned to either the image condition ($N = 102$), or the lexical condition ($N = 106$).

The experiment consisted of 24 practice trials, in which each stimulus was shown once, followed by 120 test trials, in which each

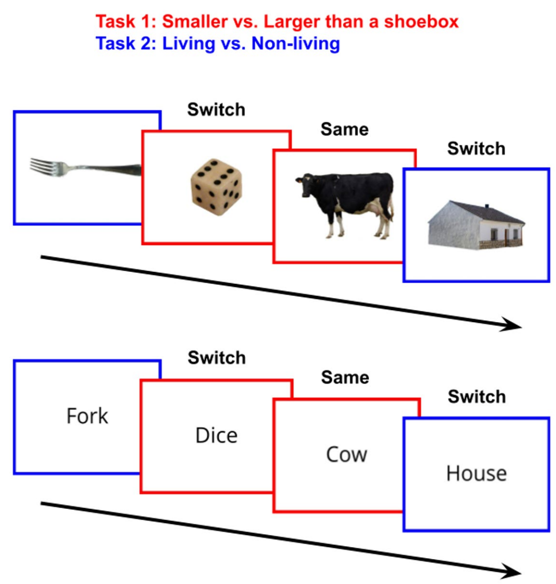

FIGURE 1
Task-switching procedure; In each trial, participants were shown an image or word stimulus surrounded by a blue or red border, which indicates the task to complete. When the task is the same as the task for the previous trial, the current trial is a "Same" trial. If it was the alternate task, the current trial is a "Switch" trial.

stimulus was shown a total of five times throughout the block, with no breaks. Each practice trial provided feedback ("Correct"/"Incorrect"), however participants were alerted after the practice block that trials would no longer provide feedback. Each stimulus was shown for 3,000 ms, during which the participant was able to respond, followed by 500 ms of feedback for the practice trials. Between each trial, a fixation cross was shown for 300 ms, with 100 ms before and after, totalling 500 ms between the end of a trial and the start of the next trial. All trials were randomized in order.

Data analysis

Language proficiency score

Each participant provided us with a self-report on the number (up to 4, in this sample) of languages with which they were familiar enough to use in social contexts. Additionally, to better reflect each participant's overall language experience, we developed a continuous total language proficiency measure of multilingualism from the rest of the LEAP-Q responses. For each component (reading, understanding, or speaking) we calculated a component score based equally on the participants' response to a series of true-false questions (the true-false score) and the self-rated proficiency score. For each of the languages spoken by the participant, a language proficiency score was calculated based on the component scores for that language, where each component was equally weighted (Figure 2). The total proficiency score is the sum of the language proficiency scores of all languages spoken by a participant. A total component proficiency score was also calculated, as the sum of the component proficiency scores of all the languages known by the participant.

While fluency in more languages provides a participant with the opportunity to generate a higher proficiency score (as reading/speaking/understanding questions are only asked about those languages that a participant initially reports familiarity with), it is not the case that number of languages perfectly predicts proficiency score, though these measures are highly correlated ($\beta = 6.14$, $R^2 = 0.822$, $p < 0.001$). While proficiency scores tend to increase as participants report familiarity with more languages, there are a number of exceptions to this general rule. For example, sometimes participants report being familiar with languages that they self-studied as a hobby, but acquired little real fluency with. Proficiency scores—which provide more detail about a participant's daily use and experience with a language—help to differentiate those participants who are using languages regularly from those who have a more passing or superficial familiarity.

Results

Working memory

Audio digit span

For all participants, the average score on the digit span task was 6.90 items (SD = 1.29). For monolingual participants, this score was 7.00 (SD = 1.23). For multilingual participants, this score was 6.87 (SD = 1.31). This difference between monolingual and multilingual participants was not significant [$t(206) = -0.62$, $p = 0.537$].

A similar result held when we treated language proficiency as a continuous variable. Neither overall proficiency [$t(206) = 0.23$, $R^2 = 0.000$, $p = 0.821$], nor any of its components [reading

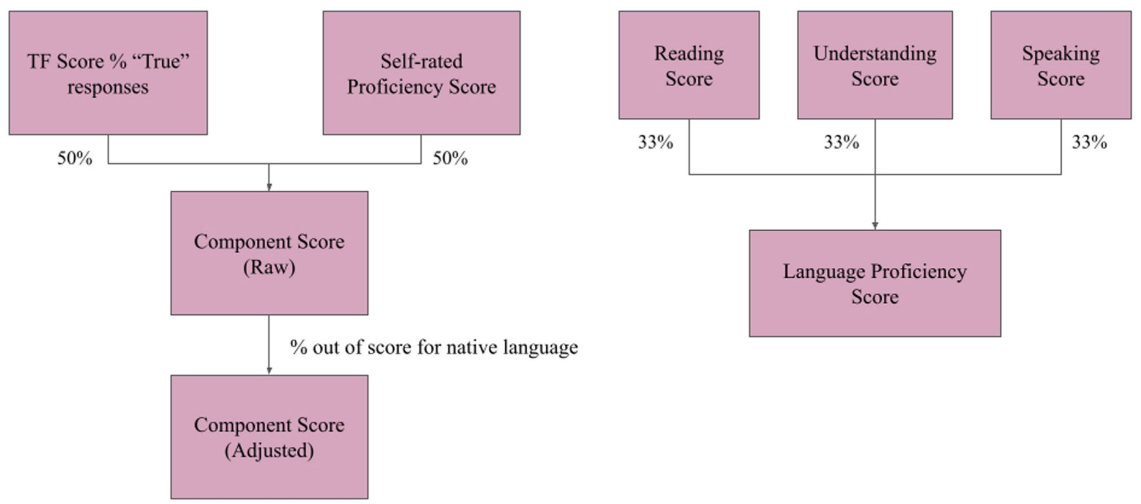

FIGURE 2
(Left) The component scores (reading, understanding, and speaking scores) were calculated by equal weighting of the True/False questionnaire responses and the self-rated proficiencies. These were normalized as a percentage out of the score for the participant's native language. (Right) Language proficiency scores for each language a participant speaks were calculated by equal weighting of each of the component scores (reading, understanding, and speaking scores).

($p = 0.889$), speaking ($p = 0.879$), understanding ($p = 0.716$)], significantly predicted audio digit span.

Visual digit span

For all participants, the average score on the visual digit span task was 6.70 items (SD = 1.46). For monolingual participants, this score was 6.30 (SD = 1.20). For multilingual participants, this score was 6.82, (SD = 1.52). This difference between monolingual and multilingual participants was significant, ($t(206) = 2.49$, $p = 0.013$).

A similar result held when we treated language proficiency as a continuous variable. A regression analysis indicated that language proficiency significantly predicted scores in the visual digit span task ($t(206) = 4.39$, $R^2 = 0.081$, $p < 0.001$). All three components were also significant predictors ($p < 0.001$).

N-back task

For all participants, we calculated the proportion of correct trials. Across all participants, the proportion of correct trials was 74.9% (SD = 15.1). The proportion of correct trials was not significantly different between monolingual participants ($M = 74.9\%$, SD = 14.5) and multilingual participants ($M = 74.9\%$, SD = 15.3; $t(206) = 0.02$, $p = 0.984$). When we treated language proficiency as a continuous variable, we found no evidence of a relationship between proficiency and the proportion of correct trials ($t(206) = 0.64$, $R^2 = 0.002$, $p = 0.526$), or between any of its components and the proportion of correct trials (all p-values are greater than 0.250).

An analysis of reaction time found no evidence of differences in performance between monolingual and multilingual participants ($t(206) = -0.34$, $p = 0.731$), or evidence of a relationship between proficiency score and reaction time ($t(206) = -0.19$, $R^2 = 0.000$, $p = 0.851$).

Summary

Consistent with prior results, our sample indicates that working memory performance is similar between monolingual and multilingual participants. The exception to this is in the visual digit span task, where participants who reported familiarity with multiple languages showed better performance than those who knew only English. It is somewhat counterintuitive that language background would have greater influence on a visual memory as opposed to auditory memory, given the auditory nature of the languages with which participants reported familiarity. As such, this result may be anomalous or spurious. It is at least inconsistent with meta-analytic work indicating that any difference in working memory tasks is most likely to be found in verbal working memory tasks (Monnier et al., 2022).

More generally, these results suggest that our monolingual sample is comparable to our multilingual sample. Even for the visual working memory task where there is a significant difference between groups, the effect size (Cohen's $d = 0.37$) is small. There is little reason to believe that these groups should show large differences in their performance on cognitive tasks. Instead, they appear to be relatively comparable, as would be expected of a group of participants who have all been selected for admission into a prestigious private university.

Word segmentation and the Simon task

Word segmentation

Across all cue conditions and pause durations, the average number of correct trials out of eight total trials in the word segmentation task was 4.18 (SD = 2.75). To assess the effect of cue conflict in word segmentation, we performed a regression analysis with cue version (Convergent or Conflicting) and pause length (0, 10, 15, 25, or 50 ms) as predictors. There was a significant main effect of cue condition [$t(198) = -6.63$, $p < 0.001$]. On average, participants' ability to identify statistically defined words (i.e., the number of "correct responses") was much higher in the Convergent condition ($M = 6.39$, SD = 1.66) than in the Conflicting condition ($M = 2.15$, SD = 1.82). This suggests that our manipulation of cue conflict was at least somewhat effective, as participants identified (statistically defined) words more easily in situations where pause cues aligned with those words than in situations where pause cues occurred in the middle of statistically-defined words in the input stream.

There was no significant main effect of pause length, nor were any pairwise comparisons between pause duration significant (all p-values greater than 0.250). For example, the participants' average performance at 0 ms ($M = 4.26$, SD = 2.51) was quite similar to their average performance at the 50 ms pause duration ($M = 5.10$, SD = 2.63). While it is difficult to interpret null effects, the lack of performance difference across pause durations is consistent with the possibility that pause cues have relatively symmetric effect. That is, they improve performance in the convergent cue languages as much as they impair performance in the conflicting cue languages, such that there is no overall effect of pause duration on performance.

As we predicted, there was an interaction between cue condition and pause length such that participants' performance was worse in the Conflicting condition ($M = 1.79$, SD = 2.04) than in the Convergent condition ($M = 6.72$, SD = 1.56) at the 50 ms pause length [$t(198) = -1.81$, $p = 0.072$], but not at any other pause lengths. This result suggests that at pause lengths of less than 50 ms, participants do not weigh the pause cue as heavily, consistent with prior results for English speakers (Weiss et al., 2010).

Language background had no apparent effect on performance in the word segmentation task, regardless of whether we considered language as a categorical factor, or treated it as a continuous variable. Across all levels of language familiarity and proficiency, there were no significant main effects of language background, nor any interactions with cue condition or pause duration.

Nevertheless, these results indicate that there is significant cue competition between statistical information and pauses at 50 ms (but not 0 ms). Therefore, we should expect that if performance in the word segmentation task is tapping into the same kind of conflict resolution processes implicated in executive function, then performance in the Simon task might be correlated with performance in the word segmentation task at 50 ms pause duration. We should not expect any correlation with performance in the 0 ms version of the word segmentation task, as there is no conflict to resolve—no pauses—in this version of the language. At pause durations between 0 and 50 ms, we might expect somewhat intermediate values of conflict resolution, but for the sake of expositional clarity, we will focus on the 0 ms condition (minimal competition to resolve) and the 50 ms condition (maximal competition to resolve) to investigate as predictors of performance in the Simon task.

Simon task

In the Simon task, participants are presented with Congruent trials, in which the presented stimulus (the word "LEFT" or the word "RIGHT") and the appropriate response are on the same side of the screen, and Incongruent trials, in which the stimulus is on the opposite side of the screen from the appropriate response. The Simon Effect is derived by comparing the average reaction time of Congruent trials to the average reaction time of Incongruent trials (which will almost always be longer, on average). A small Simon Effect means that reaction time to Incongruent trials is almost as fast as reaction time to Congruent trials, and is a sign of effective conflict resolution. A large Simon Effect means that reaction time to Incongruent trials is much slower than reaction time to Congruent trials, and is a sign of difficulty with conflict resolution.

To calculate a Simon Effect for each participant, we first removed incorrect trials from the dataset (total accuracy was as 93.90%, SD = 6.66, such that errors represented less than 7% of total trials). Of the remaining correct trials, as expected, participants were slower to respond to Incongruent trials; the average reaction time for Congruent trials was 567.70 ms (SD = 77.28), while the average reaction time for Incongruent trials was 590.72 (SD = 76.21). For each participant, we calculated a Simon Effect by subtracting their average reaction time on Congruent trials from their average reaction time on Incongruent trials.

Over all participants, the average magnitude of the Simon Effect was 23.03 ms (SD = 50.77). The magnitude of the Simon Effect was not different between monolingual and multilingual participants [$t(206) = 0.69$, $p = 0.489$], nor was it in the direction predicted by a bilingual cognitive advantage, as monolinguals had a numerically (though not significantly) smaller Simon Effect ($M = 22.7$, SD = 54.3) than their multilingual peers ($M = 23.5$, SD = 45.9). The continuous language proficiency score, on the other hand, was a marginally significant predictor of the magnitude of the Simon Effect [$t(206) = 1.68$, $R^2 = 0.013$, $p = 0.094$]. Our regression analysis indicated that higher language proficiency scores correspond to a numerically larger Simon Effect ($\beta = 1.01$). When breaking the proficiency score into components, reading proficiency [$t(206) = 1.83$, $R^2 = 0.016$, $p = 0.069$] and understanding proficiency [$t(206) = 1.68$, $R^2 = 0.013$, $p = 0.087$] were marginally significant predictors of Simon Effect magnitude, while speaking proficiency was not [$t(206) = 1.72$, $R^2 = 0.014$, $p = 0.202$].

In addition to the magnitude of the Simon Effect, we also investigated, as suggested by a thoughtful reviewer, the overall reaction time (RT) in congruent trials, in incongruent trials, and averaged over all trials. In congruent trials, the average accuracy among all participants was 96.06% (SD = 6.09) and average RT was 567.70 ms (SD = 77.28). There was no significant difference in congruent RT between monolinguals ($M = 557.63$, SD = 83.60) and multilinguals ($M = 570.88$, SD = 75.16; $t(206) = 0.99$, $p = 0.322$). A similar (lack of) relationship held when using the continuous language proficiency score to predict reaction times on congruent trials [$t(206) = 1.61$, $R^2 = 0.014$, $p = 0.108$]. Additionally, none of the components of the proficiency score significantly predicted congruent RT.

For incongruent trials, the average accuracy was 91.74% (SD = 9.59) and the average RT was 590.72 (SD = 76.21). When treating language experience as a categorical variable, there was no significant difference in incongruent RT between monolinguals ($M = 575.97$, SD = 79.79) and multilinguals [$M = 595.39$, SD = 74.69; $t(206) = 1.51$, $p = 0.132$]. However, the continuous proficiency measure was a significant predictor of incongruent RT such that greater proficiency corresponded with longer RT [$t(206) = 2.67$, $R^2 = 0.038$, $p = 0.008$]. All three components of proficiency also significantly predicted incongruent RT ($p = 0.012$ for reading, $p = 0.013$ for speaking, and $p = 0.009$ for understanding). While this is not an effect we predicted, we do note that it is at least consistent with the claim that a continuous measure of linguistic experience may be a more sensitive measure than a categorical sorting.

Across all trials, the average RT was 578.87 (SD = 72.69). The relationship with language background showed a similar pattern as incongruent RT. The overall RT was not significantly different between monolinguals ($M = 566.59$, SD = 76.88) and multilinguals [$M = 582.76$, SD = 71.12; $t(206) = 1.31$, $p = 0.192$]. At the same time, greater proficiency significantly predicted longer overall RT [$t(206) = 2.22$, $R^2 = 0.027$, $p = 0.028$], and all three of its components were also significant predictors ($p = 0.034$ for reading, $p = 0.041$ for speaking, and $p = 0.030$ for understanding). The significant effect seen while aggregating across all trials was largely driven by the difference in RT in Incongruent trials, though the (non-significant) trend in the Congruent trials appears to be in the same direction.

These results are not consistent with a multilingual advantage in conflict resolution. While categorical groupings of our participants indicated no difference in performance between monolingual and multilingual participants, our continuous measure of language background indicated that increasing language proficiency predicted a larger Simon Effect. That is, the more behavioral evidence a participant reported of using multiple languages, the slower they were to resolve conflict and produce a response in incongruent trials.

Relations between word segmentation and the Simon effect

To the extent that conflict resolution via attentional control is a domain general or task general ability, we should expect to see that participants who are good at conflict resolution in one task should also be successful in other tasks that measure conflict resolution. In particular, prior research has suggested that the Simon Effect might be related to conflict resolution in word segmentation, in both monolingual (Weiss et al., 2010) and bilingual (Bartolotti et al., 2011) populations. To determine whether we replicated this claim, we used a regression analysis to see if the Simon Effect predicted word segmentation accuracy, as a function of cue conflict and pause duration.

Our analysis indicated that the magnitude of the Simon Effect was not predictive of performance on the segmentation task. This was true when considering all participants contributing data [$t(206) = -0.54$, $R^2 = 0.001$, $p = 0.590$], the monolingual subgroup only [$t(206) = -0.73$, $R^2 = 0.005$, $p = 0.467$], or the multilingual subgroup only [$t(206) = 0.09$, $R^2 = 0.000$, $p = 0.932$; see Figure 3]. Similarly, regression analysis indicated that continuous language proficiency did not moderate the relation between Simon Effect and word segmentation performance.

As discussed above (see "Word Segmentation"), pauses become a more powerful cue as their duration increases, and at 50 ms there appears to be maximum conflict between pause cues and statistical cues. By contrast, at 0 ms, there is minimal conflict, because pauses (definitionally) do not occur. Therefore, we assessed whether pause duration moderated the relationship between Simon score and word segmentation task. Our analyses indicated that there was no systematic relationship between these variables (see Figure 4). Indeed, the relationship between Simon task and word segmentation was stronger at 0 ms (where there is no possible pause-related conflict resolution in the word segmentation task) than at 50 ms.

Summary

These results provide little evidence to support claims for a multilingual advantage. Our primary hypothesis, that conflict resolution as measured in the Simon task would be predictive of conflict resolution in a linguistic task—word segmentation in the face of conflicting cues—was not supported. This fails to replicate prior work observing such a relationship in both monolingual (Weiss et al., 2010) and multilingual (Bartolotti et al., 2011) populations, so our failure to find evidence of the relation in either group is noteworthy.

One possibility that may explain our failure is that while our manipulation of pause duration (50 ms) was identical to that used in past studies, this pause duration may have been more (or less) salient to our participants than has been the case in prior studies. In particular, prior studies assessing the relation between segmentation performance and Simon Effect have used segmentation languages that are more complex than ours in a variety of ways, such as having more words, longer words, or more challenging perceptual features than the relatively simple 4 bisyllabic word language used in this study.

At the same time as we failed to find support for our primary hypothesis, we also found evidence that was directly contradictory

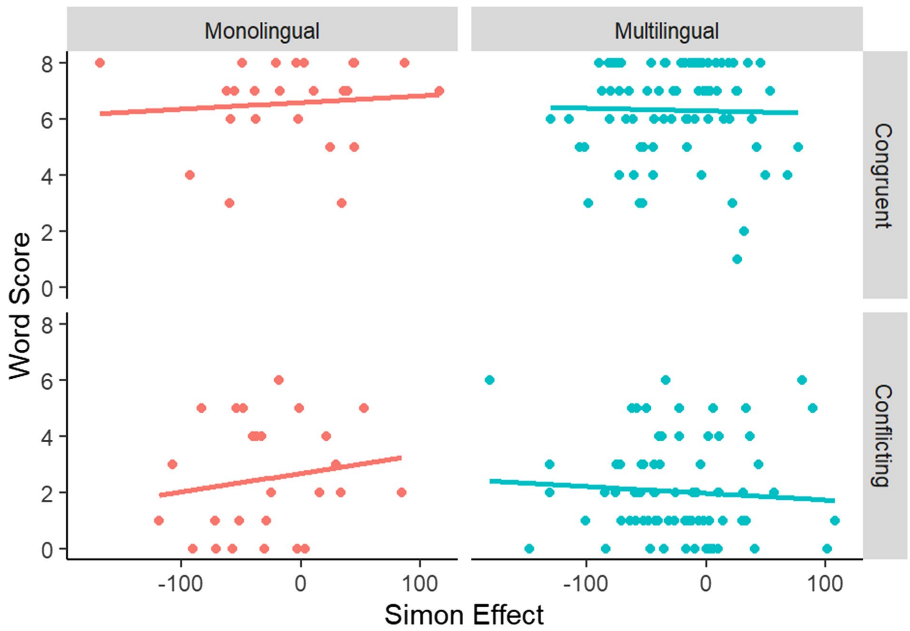

FIGURE 3
Relationship between the Simon Effect (ms) and performance on the Word Segmentation task (out of 8 questions), by multilingualism and word segmentation cue version.

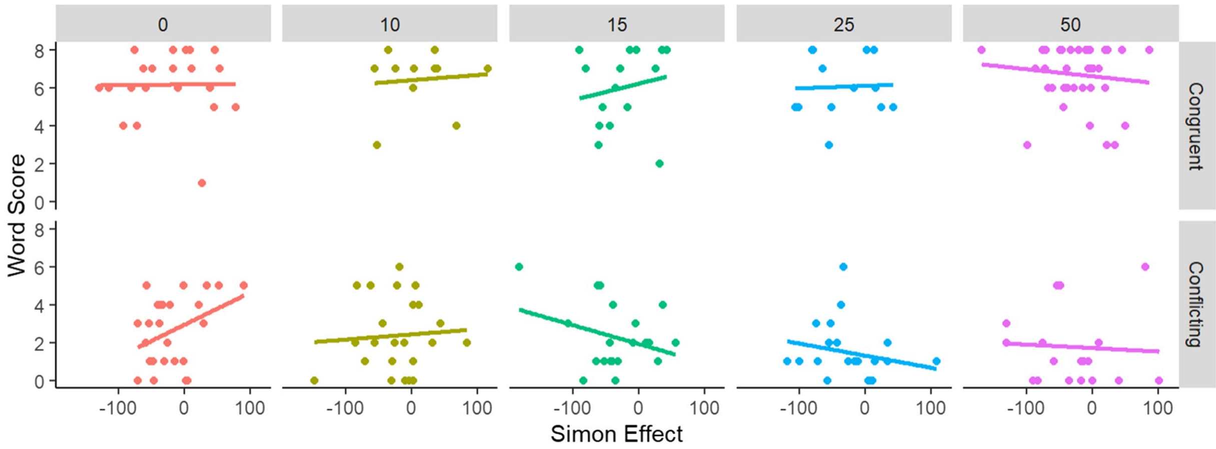

FIGURE 4
Relationship between the Simon Effect (ms) and performance on the Word Segmentation task (out of 8 questions), by pause length and word segmentation cue version.

to claims of a multilingual advantage: multilingual participants appeared to perform worse in the Simon task. This is inconsistent with several prior published results suggesting a bilingual or multilingual advantage for conflict resolution (e.g., Bialystok et al., 2004). These results are not totally unprecedented, as not all investigators have found a multilingual advantage in conflict resolution (e.g., Morton and Harper, 2007). Moreover, it is worth noting that our results indicating a "monolingual advantage" were only significant when we treated language background as a continuous measure.

Dimension switching

In the Dimension-Switching task, participants are presented with a stimulus (either a word, or an image), and instructed by a colored border around the stimulus to make a particular judgment (living/non-living or large/small) about the stimulus. On trials where participants are making a different judgment than the one they made on the prior trial, their reaction should be somewhat slower than if they are repeating the same judgment from the previous trial. This "switch cost" is the primary dependent variable of interest in the task. Prior work suggests that multilingual speakers should have a smaller switch cost than monolingual speakers (e.g., Prior and MacWhinney, 2010).

Performance was similar across different language backgrounds. Multilingual participants had numerically smaller switch costs ($M = 284.33$, $SD = 163.91$) compared to monolingual participants ($M = 299.06$, $SD = 160.00$; see panel A in Figure 5). Consistent with this observation, regression analysis indicated that participants with higher language proficiency score had lower switch cost ($\beta = -1.31$; see panel A in Figure 6). However, neither the monolingual-multilingual dichotomy [$t(204) = -1.27$ $p = 0.205$] nor the continuous language proficiency [$t(204) = -1.52$, $p = 0.131$] significantly predicted switch cost.

Dimension-switching and linguistic stimuli

Of interest to us is whether switch cost is more closely related to language background when the dimension-switching task uses explicitly linguistic stimuli (orthographic words) or when it uses stimuli that are less directly tied to linguistic representations (photorealistic visual images). We instantiated this modality difference as a between-subjects variable. Overall performance was very similar between the orthographic word and photographic images versions of the task (see Figure 7). Across the two versions of the task, there is no overall difference in accuracy [$t(206) = 0.94$, $p = 0.350$] or reaction time [$t(206) = -0.73$, $p = 0.464$]. Note, however, that there is some hint of a difference in switch costs across modality, as switch costs were somewhat greater with lexical stimuli, a difference that was marginally significant [$t(206) = 1.76$, $p = 0.079$].

To assess whether language background was differentially related to switch cost as a function of stimulus modality, we used linear regression with three predictors: language background, stimulus modality, and the interaction term. We fit two models, one treating language background as a binary classification (monolingual, multilingual) and one using the continuous proficiency score. Stimulus modality did not emerge as a significant predictor, regardless of whether language background is included as a categorical variable [$t(204) = -0.23$, $p = 0.816$] or a continuous variable [$t(204) = -0.90$, $p = 0.367$]. This null effect is consistent with the similar performance across the pictographic and orthographic versions of the task described above.

The same regression analyses indicated no significant interaction between stimulus modality and language background. In the image version of the task, the switch cost was numerically smaller for multilingual participants ($M = 255.06$, $SD = 156.73$) than for monolingual participants ($M = 304.24$, $SD = 170.73$); in the word version, the switch cost was higher for multilingual participants ($M = 311.46$, $SD = 166.68$) than for monolingual participants ($M = 293.44$, $SD = 166.68$; see panel B in Figure 5). Consistent with these descriptions of the data, our regression

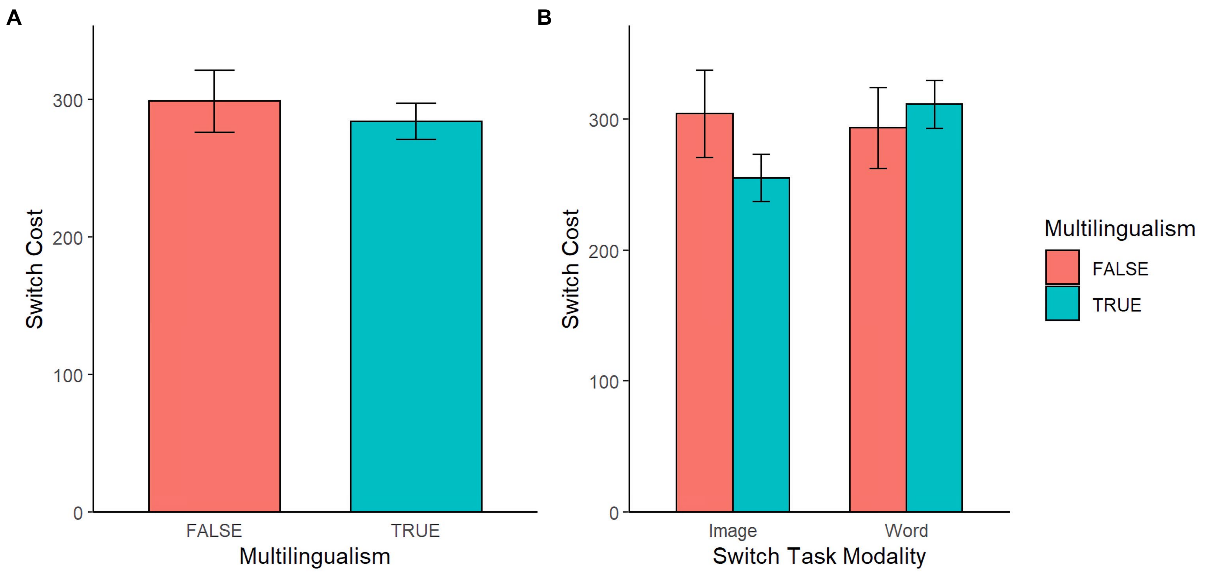

FIGURE 5
(A) Switch task performance by multilingualism; and **(B)** switch task performance by multilingualism and switch task modality.

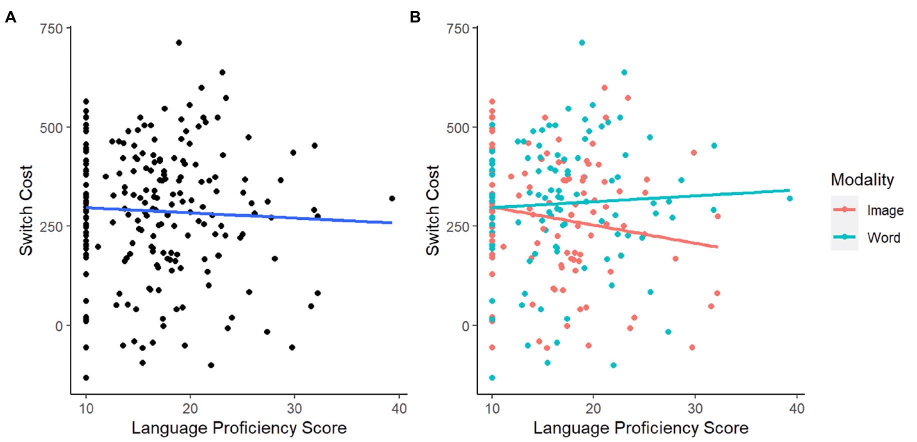

FIGURE 6
(A) Switch task performance by language proficiency; **(B)** switch task performance by language proficiency and switch task modality.

analysis showed that language background predicted lower switch cost in the image version ($\beta = -4.56$), and predicted higher switch cost in the word version ($\beta = 1.47$; see panel B in Figure 6). However, these interactions between stimulus modality and language background were not found to be significant in either model [$t(204) = 1.26$, $p = 0.208$ with categorical, $t(204) = 1.57$, $p = 0.119$ with continuous].

Dimension-switching and word segmentation

As we noted in the introduction, transfer across tasks is more likely when those tasks are similar to some extent. Because the dimension-switching task relied upon the linguistic nature of the stimuli, we were prompted by a helpful comment from reviewers to investigate the relation between the dimension-switching task and word segmentation task, which is also putatively linguistic in nature. Switch cost was not a significant predictor of performance on the word segmentation task [$t(206) = 0.93$, $R^2 = 0.004$, $p = 0.355$]. When including language experience in the model, none of the main effects or interactions were significant (all p-values greater than 0.200), regardless of whether the binary classification of monolinguals and multilinguals or the continuous language proficiency score (composite or any of the components) was used.

Summary

As with the relationship between Simon Effect and word segmentation, our investigation of the relationship between

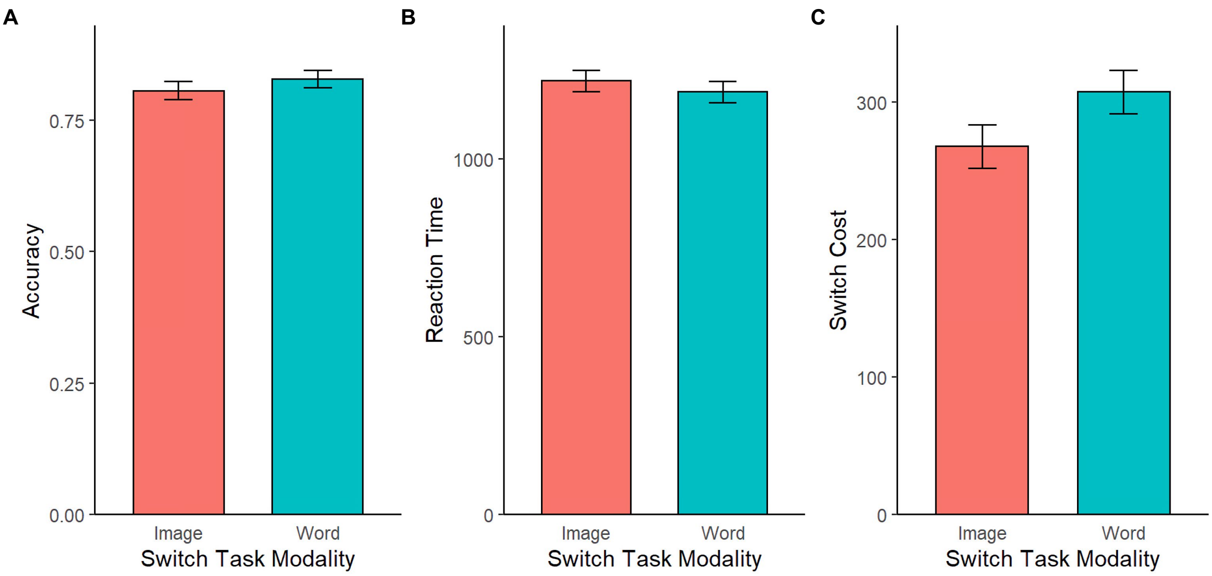

FIGURE 7
Effect of switch task modality on accuracy **(A)**, reaction time **(B)**, and switch cost **(C)**. There were no significant differences in accuracy, reaction time, or switch cost across the two task versions.

dimension switching and stimulus modality did not support our hypotheses. We did not find that language background more strongly predicted performance with linguistic stimuli than performance with pictorial stimuli. Indeed, while none of our effects were significant, the relationship that seems most robust is between linguistic background and performance in a pictorial version of the task, where multilingual speakers show most robust evidence of an advantage with pictorial stimuli, while monolingual participants appear to have an advantage for orthographic stimuli. These results, while potentially suggestive of patterns that might be revealed by a larger sample or more sensitive design, are the opposite of what we predicted. Similarly, we failed to replicate prior results that multilingual participants have an advantage in dimension switching (e.g., Prior and MacWhinney, 2010; Wiseheart et al., 2016).

Language dominance

The primary novel hypothesis that motivated this research (in addition to the desire to attempt a conceptual replication of Weiss et al., 2010) was that the "distance of transfer" would predict the degree to which multilingual participants show a difference in executive function skills from monolingual participants. That is, we predicted that monolingual and multilingual participants would perform in a relatively similar fashion for non-linguistic tasks, and less similarly in linguistic tasks, a prediction we attempted to directly assess with our manipulation of cue presentation in the task switching task.

However, our design may work against our ability to detect such differences. Specifically, we conflated the "linguistic" nature of our stimuli with the English language. It might be the case that multilingual speakers are indeed more successful in executive function tasks that involve linguistic stimuli or processes. But because all of our linguistic stimuli are instantiated in English, our multilingual participants' advantage may be masked by a general advantage for monolingual speakers when tested in their native language.

As suggested during the review process, we sought to assess the plausibility of this alternative explanation by more thoroughly investigating how participants' self-reported language dominance (that is, the language with which they feel most comfortable, fluent, and practiced) was related to their task performance. To the extent that a multilingual advantage is masked by a general multilingual disadvantage (compared to monolingual speakers) with English stimuli, we might expect that this disadvantage would be less severe for those multilingual participants who report that English is their dominant language. Conversely, if multilingual participants show relatively similar performance, regardless of their language dominance, this would suggest that an English-mask of the multilingual advantage is less plausible (though certainly not impossible). Additionally, a more extensive investigation of language background and language dominance may provide us with some insight about the generalizability of our results.

As such, the analyses below investigate performance in our 158 multilingual participants, sorted by whether they report English as their dominant language (120 participants), or report that some other language (38 participants) is dominant

10.3389/fpsyg.2022.923123

in their daily use. We explored whether this sorting predicts differences in the Simon task, which is the task for which we have the most data for each participant, and thus presumably the task most sensitive to individual differences. Additionally, we analyzed data from the Dimension Switching task, which we explicitly designed to test the effect of linguistic stimuli.

Simon task

In terms of accuracy, the 120 English-dominant participants ($M = 0.937$, $SD = 0.072$) and the 38 Other-dominant participants ($M = 0.938$, $SD = 0.067$) showed similar performances ($p = 0.984$). This was also true when analyzing only congruent ($p = 0.867$) or only incongruent trials ($p = 0.888$).

Overall reaction times (RT) of English-dominant participants ($M = 578.18$, $SD = 72.07$) were not significantly different ($p = 0.14$) than the reaction times of Other-dominant participants ($M = 597.22$, $SD = 66.88$). A linear regression with multilingualism as a continuous variable also showed that differences between the English-dominant and Other-dominant participants were not significant ($p = 0.388$). These results held for analyses within congruent and incongruent trials, such that both t-tests and linear regressions incorporating the multilingualism continuous variable resulted in non-significant effects of language-dominance on RT (all $p > 0.10$). The only marginally significant relation found was between multilingualism as a continuous score and RT specifically in incongruent trials [$t(154) = 1.82$, $p = 0.070$]. This relationship was not significant in congruent trials and over all trials averaged. Note that these exploratory analyses are not corrected for multiple comparisons (to allow for greatest sensitivity to possible patterns of interest), so marginal significance and even significant effects should be interpreted cautiously.

English-dominant participants showed a numerically smaller Simon Effect ($M = 24.28$, $SD = 48.59$) than Other-dominant participants ($M = 25.24$, $SD = 51.29$). However, this difference was not significant ($p = 0.920$). A linear regression with multilingualism as a continuous variable also indicated non-significant effects ($p = 0.409$).

Dimension switching

English-dominant participants had a mean accuracy of 0.799 ($SD = 0.181$), and Other-dominant participants had a mean accuracy of 0.849 ($SD = 0.158$). This difference was not significant ($p = 0.110$). Analysis of accuracy in the image version of the task between English-dominant ($N = 59$) participants ($M = 0.798$, $SD = 0.172$) and Other-dominant ($N = 17$) participants ($M = 0.796$, $SD = 0.191$) also showed no significant differences ($p = 0.966$).

However, in the word version of the task, Other-dominant participants ($N = 21$) showed higher accuracy ($M = 0.892$, $SD = 0.112$) than English-dominant ($N = 61$) participants ($M = 0.800$, $SD = 0.192$). This effect is significant [$t(80) = 2.59$, $p = 0.011$], but counterintuitive, as we would not expect English-dominance to be associated with lower accuracy on a task presented in English. The effect is also potentially consistent with a speed-accuracy tradeoff, meaning that this result cannot be interpreted in isolation from data about reaction time.

Overall, reaction times (RT) of English-dominant participants ($M = 1187.50$, $SD = 323.15$, $N = 120$) and those of Other-dominant participants ($M = 1233.20$, $SD = 305.44$, $N = 38$) did not significantly differ ($p = 0.434$). Upon analyzing differences between the English-dominant and non-English dominant groups within the image and word task modalities, we found that there were no significant differences in either modality ($p = 0.979$ for image version, $p = 0.287$ for word version). The lack of differences in reaction time is not consistent with the possibility that a speed-accuracy tradeoff explains the Other-dominant participants' surprising advantage in accuracy on the lexical version of the task (though the lack of a significant effect certainly does not *rule out* the possibility of a speed-accuracy tradeoff).

Lastly, we analyzed differences in switch cost across participants. Recall that the switch cost is the primary dependent variable of interest in the Dimension-Switching paradigm; a low switch cost indicates stronger executive function, while a high switch cost indicates difficulty with the demands of the task. Overall, there was no significant difference ($p = 0.145$) in the switch cost between English-dominant participants ($M = 273.19$, $SD = 161.07$, $N = 120$) and Other-dominant participants ($M = 319.52$, $SD = 169.96$, $N = 38$). Similarly, the switch costs in the image version of the task showed no significant difference ($p = 0.715$) between the English-dominant participants ($M = 250.84$, $SD = 146.97$, $N = 59$) and the Other-dominant participants ($M = 269.71$, $SD = 191.20$, $N = 17$).

However, there was a marginally significant difference found in the lexical version of the task as a function of language dominance [$t(80) = 1.67$, $p = 0.098$]. Here, English-dominant participants ($M = 294.81$, $SD = 172.08$, $N = 61$) showed lower switch costs than Other-dominant participants ($M = 359.84$, $SD = 142.74$, $N = 21$). One could interpret this result as partially consistent with the hypothesis that (a subset of) multilingual speakers show a selective advantage on lexically instantiated executive function tasks. But even here, it should be noted that our English-dominant multilinguals' switch cost on the task is virtually identical to that of monolingual participants ($M = 293.44$, $SD = 166.68$) in the lexical version of the dimension-switching task. Therefore, these results may indicate a selective disadvantage for Other-dominant multilingual participants in the lexical version of the task switching task. This result could suggest that engaging executive function is more challenging in non-dominant language contexts.

Summary

In the Simon task, our results indicate that participants' language dominance had little effect on performance. This is consistent with the more general lack of evidence across our tasks, with the hypothesis that language background is related to performance in our executive function tasks. Interestingly, in the Dimension-Switching task, language dominance was related to both accuracy and the magnitude of switch cost. Due to the exploratory nature of these analyses, these results should be interpreted cautiously. But they are consistent with the claim that some degree of executive function performance may be masked or inhibited by the (English) linguistic nature of some of our stimuli.

Discussion

At some level, it is inarguable that experience with multiple languages shapes the cognitive system of the people who speak them. An English speaker automatically connects the word "cat" to their representation of a whiskered mammal that purrs. An English-Spanish bilingual also does so for the word "gato," in a way that the monolingual English speaker does not. The learning challenges and processes in a multilingual environment are different than those in a monolingual environment (e.g., Byers-Heinlein and Fennell, 2014; Singh, 2021). As such, our question is not "does experience with multiple languages change cognition," but rather, what is the extent of those changes? The minimalist stance, consistent with a modular view of cognition (Fodor, 1983), is that experience with language only influences linguistic processing (e.g., Goldsmith and Morton, 2018; Dick et al., 2019). The maximalist view is that lifelong experience with multilingualism provides a domain-general strengthening of attentional processes associated with executive function, especially selection of relevant information, suppression of irrelevant information, and conflict resolution (e.g., Craik et al., 2010; Prior and MacWhinney, 2010; Bialystok, 2017). Or, to put it in terms borrowed from cognitive science, the central question is whether language experience "transfers" to only relatively close tasks, as is often seen with laboratory investigations of learning (e.g., Sala and Gobet, 2017). The alternative possibility is that language—perhaps due to our extensive experience with it, or its centrality to cognition—serves as a basis for far transfer, and has an influence that can be felt in a wide variety of different tasks, even those that are only minimally related to language (except by virtue of sharing some common underlying process).

Recent surveys of the field and meta-analyses (e.g., Donnelly et al., 2019; Vīnerte and Sabourin, 2019; Gunnerud et al., 2020; Paap et al., 2020; de Bruin et al., 2021) make a compelling case that both the maximalist position—of essentially unlimited far transfer to a variety of executive function tasks—and the minimalist position—expecting no transfer outside of linguistic tasks—are all but unsustainable in their purest forms. Rather than a consistent pattern of success (or failure) in transfer to cognitive tasks, the literature presents us a "mixed bag" of findings, with some converging replications of a multilingual advantage (e.g., Anderson et al., 2020a; Brini et al., 2020), several inconsistent or small effects (e.g., Lukasik et al., 2018; Paap et al., 2018), and even contradictory evidence of monolingual advantages in cognitive performance (e.g., Paap and Greenberg, 2013; Nichols et al., 2020). This presents us with the challenge of determining, not whether or not a multilingual advantage exists, but what are the modulating factors or constraints that determine when differences will exist across populations (van den Noort et al., 2019). This determination, however, is challenging given that the heterogeneity in multilingual participants and cognitive assessments provides a multitude of degrees of freedom that make falsification challenging (for discussion, see Struys et al., 2018). In such a situation, the value of theoretical frameworks as sources of falsifiable predictions is especially high (for discussion, see de Bruin et al., 2021).

One such theoretical perspective that has been advanced is a neuroscientific one (e.g., Vīnerte and Sabourin, 2019). This should be distinguished from the use of brain-based dependent variables, such as ERP or FMRI. Perhaps unsurprisingly, research efforts with these neural measures have produced the same type of mixed results as research efforts with more traditional behavioral measures (e.g., Leivada et al., 2021). Rather, we use the term neuroscientific theoretical perspective to mean research endeavors whose hypotheses are informed by an understanding of the structural, functional, and network characteristics of the brain (e.g., Hernandez et al., 2019; Del Maschio et al., 2020). For example, the Adaptive Control Hypothesis (Green and Abutalebi, 2013; Abutalebi and Green, 2016) generates predictions about the brain regions and cognitive processes that should be impacted by multilingual experience as a function of the dimensions on which those languages overlap, complement, and interfere with each other, and how those dimensions relate to known neural networks. While the Adaptive Control Hypothesis and other neurally inspired frameworks have inspired much productive research, a preliminary conclusion is that these theories are also incomplete, or fail to explain aspects of the data (for discussion, see Kałamała et al., 2020; Paap et al., 2021).

We propose that an embodiment perspective may enrich or complement the neuroscientific perspective, and other perspectives, and generate explanations and predictions about the relation between linguistic experience and cognitive processes. In particular, we hypothesized that multilingual participants would be more likely to show transfer from linguistic experience to executive control tasks (and thus, an advantage over monolingual participants) when the tasks prompted participants to engage in the same kinds of representations or processes that they engage in naturalistic language use. From this perspective, one reason why the prior literature on multilingual differences in executive control is mixed is that some tasks are much more likely to incorporate or evoke linguistic stimuli, which should make them more likely to detect differences between multilingual and monolingual participants.

Admittedly, the data that we generated to investigate predictions arising from an embodiment perspective are not tremendously consistent with our hypotheses. As noted in the results section, we repeatedly failed to support the hypothesis of a multilingual advantage in executive function—and indeed, even found evidence of a monolingual advantage in the Simon task—nor did we find evidence consistent with the claim that the linguistic nature of a task predicts the degree of multilingual advantage. However, we would like to suggest that these results should not be read as an indictment of the embodiment perspective. Instead, we believe that on further reflection, these results reflect essential limitations of our methodology, limitations that can be ameliorated by a more thorough integration of the embodiment perspective into this work.

First, some of the tasks that we chose to use may not have been sensitive to the individual differences between participants that we hoped to detect. In particular, the word segmentation task is likely to be a poor source of information about individual differences (e.g., Erickson et al., 2016; Siegelman et al., 2017a). This is not to say that the word segmentation task is uninformative. However, it is plausibly the case that the word segmentation task is informative about the differences between groups (such as monolingual and multilingual speakers) but not sensitive to individual differences among participants (for discussion, see Siegelman et al., 2017b). Because our design was aimed at assessing the relation between an individual's Simon Effect and that same individual's performance in the word segmentation task, we are focused on a level of analysis where the word segmentation task may be minimally sensitive.

A second limitation of our methodology is that we recruited a set of participants who are fairly heterogeneous in their language background and use, and at the same time—by virtue of their selection into an academically rigorous private university—not representative of the range of variation in cognitive performance and educational background seen in the population at large. In particular, while we found evidence that treating language background as a continuous variable can be informative, it is likely the case that our analyses overlook some of the continuous dimensions that are likely to shape the cognitive impact of language use, in particular age of acquisition and the degree of overlap or similarity between the languages. From an embodiment perspective, these questions are crucial, as they shape the way participants approach the task of language acquisition and use, which—in turn—should shape the kinds of tasks to which they transfer that experience. Indeed, the very heterogeneity of language backgrounds, and potential differences in participant populations, has led some to contemplate the possibility that the research question may not be tenable (e.g., Paap et al., 2014; Hartsuiker, 2015, though see Leivada et al., 2021). At the very least, it seems clear that investigations with even larger (and more representative) samples than what we collected in the current research may be necessary to assess some of these predictions about language background.

Relatedly, it is worth emphasizing our theoretical claim that embodiment theory makes predictions about when we should (and should not) see effects of language background on cognitive tasks. That is, it proposes that multilingual speakers should have an advantage in tasks that take advantage of linguistic processes and representations. Conversely, seeing a difference between monolingual and multilingual speakers should be less likely when the task in question does not evoke prior linguistic experiences. Fundamentally, we can only examine the claim—that the use of linguistic materials makes the observation of the multilingual advantage more likely—if we can observe the multilingual advantage in the first place. The fact that we failed to observe such an advantage means that we should be especially cautious about claiming which factors might or might not moderate such an effect. However, we note that our analysis of language dominance, suggested by a reviewer, is perhaps indicative of some effect of linguistic background that our design lacked the sensitivity or the sample to detect.

Finally, and perhaps most importantly, our attempt to differentiate "linguistic" stimuli from "non-linguistic" stimuli suffers from a confound. In all cases, our linguistic materials were presented in English, which is (necessarily) the primary and dominant language for our monolingual speakers, while our multilingual speakers varied more widely in their familiarity with and use of the English language. This may explain why we found an unexpected monolingual advantage in conflict resolution in the Simon task, as the conflict was always expressed with English words. To better understand differences in performance between monolingual and multilingual participants, it may be necessary—as suggested by a thoughtful reviewer—to assess performance in a non-linguistic version of the task.

Indeed, the very embodiment perspective we have advocated suggests that we miss an important avenue toward explanatory power when we classify participants by abstract terms such as "monolingual" or "multilingual," even when those terms are quantified in relatively continuous ways. Beyond this, we need to think about participants' goals, their expectations, and how their prior experiences relate to specific tasks. Doing so, we believe, will enable us to make more confident predictions about when prior linguistic experience will transfer to a task, and when it will not. While our current results do not demonstrate this principle as effectively as we hoped, we do believe that these results, in conversation with some of the other topics raised in this special issue, may point toward a useful path forward. Repeatedly, we found that where there was a "bilingual advantage," it was for stimuli that were less determinedly linguistic (images rather than words; visual digits rather than spoken digits in our span task). In debriefing afterward, however, several participants told us that these "less" linguistic stimuli actually resulted in the participant doing more linguistic processing (trying to self-generate a label). Or, to put it more broadly, a task is linguistic not because of the stimuli in the task, but because of how the participant perceives, represents, and manipulates those stimuli. The embodiment perspective, with its focus on exactly this level of analysis, may provide a productive avenue of defining the cognitive processes associated with language, and thus making predictions about which tasks should show effects of linguistic experience.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Carnegie Mellon University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

ET was the primary investigator in the lab where this work occurred, and helped to direct implementation. YK did important conceptual and pragmatic work to bring the task switching methodology into the paradigm. ZY was responsible for the final set of analyses, building upon work by ZL and W-JW. All authors contributed to the article and approved the submitted version.

Funding

The publication of this work is supported by the Carnegie Mellon University library system.

Acknowledgments

We thank Maureen Hilton for her assistance with study administration and Alice Russell for her assistance in data collection and cleaning.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.923123/full#supplementary-material

References

Abutalebi, J., and Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Biling. Lang. Congn.* 19, 689–698. doi: 10.1017/S1366728916000225

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803

Anderson, J. A., Hawrylewicz, K., and Bialystok, E. (2020b). Who is bilingual? Snapshots across the lifespan. *Biling. Lang. Congn.* 23, 929–937. doi: 10.1017/S1366728918000950

Anderson, J. A., Hawrylewicz, K., and Grundy, J. G. (2020a). Does bilingualism protect against dementia? A meta-analysis. *Psychon. Bull. Rev.* 27, 952–965. doi: 10.3758/s13423-020-01736-5

Anderson, J. A., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9

Antón, E., Carreiras, M., and Duñabeitia, J. A. (2019). The impact of bilingualism on executive functions and working memory in young adults. *PLoS One* 14:e0206770. doi: 10.1371/journal.pone.0206770

Antón, E., García, Y. F., Carreiras, M., and Duñabeitia, J. A. (2016). Does bilingualism shape inhibitory control in the elderly? *J. Mem. Lang.* 90, 147–160. doi: 10.1016/j.jml.2016.04.007

Aslin, R. N., Saffran, J. R., and Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9, 321–324. doi: 10.1111/1467-9280.00063

Barsalou, L. W. (1999a). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/S0140525X99002149

Barsalou, L.W. (1999b). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes* 28:1, 61-80, doi: 10.1080/01638539909545069

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi: 10.3389/fpsyg.2011.00324

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Biling. Lang. Congn.* 12, 3–11. doi: 10.1017/S1366728908003477

Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099

Bialystok, E. (2021). Bilingualism: pathway to cognitive reserve. *Trends Cogn. Sci.* 25, 355–364. doi: 10.1016/j.tics.2021.02.003

Bialystok, E., Craik, F. I., and Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45, 459–464. doi: 10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290

Blumenfeld, H. K., and Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *J. Cogn. Psychol.* 25, 547–567. doi: 10.1080/20445911.2013.812093

Brini, S., Sohrabi, H. R., Hebert, J. J., Forrest, M. R., Laine, M., Hämäläinen, H., et al. (2020). Bilingualism is associated with a delayed onset of dementia but not with a lower risk of developing it: a systematic review with meta-analyses. *Neuropsychol. Rev.* 30, 1–24. doi: 10.1007/s11065-020-09426-8

Byers-Heinlein, K., and Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: insights from bilingual infants. *Dev. Psychobiol.* 56, 274–291. doi: 10.1002/dev.21167

Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., et al. (2018). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat. Rev. Neurosci.* 19, 701–710. doi: 10.1038/s41583-018-0068-2

Calabria, M., Hernández, M., Branzi, F. M., and Costa, A. (2012). Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Front. Psychol.* 2:399. doi: 10.3389/fpsyg.2011.00399

Chomsky, N. (1964). [the development of grammar in child language]: discussion. *Monogr. Soc. Res. Child Dev.* 29, 35–42. doi: 10.2307/1165753

Cole, R. A., and Jakimik, J. (1980). How are syllables used to recognize words? *J. Acoust. Soc. Am.* 67, 965–970. doi: 10.1121/1.383939

Craik, F. I., Bialystok, E., and Freedman, M. (2010). Delaying the onset of Alzheimer disease: bilingualism as a form of cognitive reserve. *Neurology* 75, 1726–1729. doi: 10.1212/WNL.0b013e3181fc2a1c

De Bruin, A., Dick, A. S., and Carreiras, M. (2021). Clear theories are needed to interpret differences: perspectives on the bilingual advantage debate. *Neurobiol. Lang.* 2, 433–451. doi: 10.1162/nol_a_00038

Del Maschio, N., Sulpizio, S., and Abutalebi, J. (2020). Thinking outside the box: the brain-bilingualism relationship in the light of early neurobiological variability. *Brain Lang.* 211:104879. doi: 10.1016/j.bandl.2020.104879

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., et al. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nat. Hum. Behav.* 3, 692–701. doi: 10.1038/s41562-019-0609-3

Donnelly, S., Brooks, P. J., and Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychon. Bull. Rev.* 26, 1122–1147. doi: 10.3758/s13423-019-01567-z

Erickson, L. C., Kaschak, M. P., Thiessen, E. D., and Berry, C. A. (2016). Individual differences in statistical learning: conceptual and measurement issues. *Collabra* 2:14. doi: 10.1525/collabra.41

Fodor, J. (1983). *The Modularity of Mind*. Cambridge, Bradford Books: The MIT Press.

Foy, J. G., and Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *Int. J. Biling.* 18, 717–729. doi: 10.1177/1367006912472263

Freeman, M. R., Blumenfeld, H. K., and Marian, V. (2017). Cross-linguistic phonotactic competition and cognitive control in bilinguals. *J. Cogn. Psychol.* 29, 783–794. doi: 10.1080/20445911.2017.1321553

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas-Guasch, N., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011

Gobet, F., and Sala, G. (2020). Cognitive training: A field in search of a phenomenon. doi: 10.31234/osf.io/vxzq9

Goldsmith, S. F., and Morton, J. B. (2018). Time to disengage from the bilingual advantage hypothesis. *Cognition* 170, 328–329. doi: 10.1016/j.cognition.2017.09.011

Goodman, N. (1972). "Seven strictures on similarity" in *Problems and Projects*. Indianapolis, IN: Bobbs-Merrill.

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Gunnerud, H. L., Ten Braak, D., Reikerås, E. K. L., Donolato, E., and Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychol. Bull.* 146, 1059–1083. doi: 10.1037/bul0000301

Hartsuiker, R. J. (2015). Why it is pointless to ask under which specific circumstances the bilingual advantage occurs 73336–337. *Cortex* 73, 336–337. doi: 10.1016/j.cortex.2015.07.018

Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., Castilla-Earls, A. P., Sun, L., Weiss, S. D., et al. (2019). Neuroemergentism: a framework for studying cognition and the brain. *J. Neurolinguistics* 49, 214–223. doi: 10.1016/j.jneuroling.2017.12.010

Johnson, E. K., and Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* 44, 548–567. doi: 10.1006/jmla.2000.2755

Jusczyk, P. W., Cutler, A., and Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development* 64, 675–687.

Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., and Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: a test of the adaptive control hypothesis. *Cognition* 204:104373. doi: 10.1016/j.cognition.2020.104373

Kassai, R., Futo, J., Demetrovics, Z., and Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near-and far-transfer effects among children's executive function skills. *Psychol. Bull.* 145, 165–188. doi: 10.1037/bul0000180

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., De Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142

Leivada, E., Westergaard, M., Duñabeitia, J. A., and Rothman, J. (2021). On the phantom-like appearance of bilingualism effects on neurocognition:(how) should we proceed? *Biling. Lang. Congn.* 24, 197–210. doi: 10.1017/S1366728920000358

Lukasik, K. M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., and Laine, M. (2018). Bilingualism and working memory performance: evidence from a large-scale online study. *PLoS One* 13:e0205916. doi: 10.1371/journal.pone.0205916

Macnamara, B. N., and Conway, A. R. (2014). Novel evidence in support of the bilingual advantage: influences of task demands and experience on cognitive control and working memory. *Psychon. Bull. Rev.* 21, 520–525. doi: 10.3758/s13423-013-0524-y

Marian, V., Blumenfeld, H.K., and Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *JSLHR* 50, 940–967. doi: 10.1044/1092-4388(2007/067)

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Monnier, C., Boiché, J., Armandon, P., Baudoin, S., and Bellocchi, S. (2022). Is bilingualism associated with better working memory capacity? A meta-analysis. *Int. J. Biling. Educ. Biling.* 25, 2229–2255. doi: 10.1080/13670050.2021.1908220

Moreno-Martínez, F. J., and Montoro, P. R. (2012). An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS One* 7:e37527. doi: 10.1371/journal.pone.0037527

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., and Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *J. Child Lang.* 41, 600–633. doi: 10.1017/S0305000913000111

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., and Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: a population study of executive function in 11,000 people. *Psychol. Sci.* 31, 548–567. doi: 10.1177/0956797620903113

Paap, K. R., Anders-Jefferson, R., Mason, L., Alvarado, K., and Zimiga, B. (2018). Bilingual advantages in inhibition or selective attention: more challenges. *Front. Psychol.* 9:1409. doi: 10.3389/fpsyg.2018.01409

Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002

Paap, K. R., Mason, L., and Anders-Jefferson, R. (2021). Predictions about the cognitive consequences of language switching on executive functioning inspired by the adaptive control hypothesis fail more often than not. *Brain Sci.* 11:1217. doi: 10.3390/brainsci11091217

Paap, K. R., Mason, L., Zimiga, B., Ayala-Silva, Y., and Frost, M. (2020). The alchemy of confirmation bias transmutes expectations into bilingual advantages: a tale of two new meta-analyses. *Q. J. Exp. Psychol.* 73, 1290–1299. doi: 10.1177/1747021819900098

Paap, K. R., Sawi, O. M., Dalibar, C., Darrow, J., and Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be extraordinarily difficult to discover. *AIMS Neurosci.* 1, 245–256. doi: 10.3934/Neuroscience.2014.3.245

Peal, E., and Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychol. Monogr. Gen. Appl.* 76, 1–23. doi: 10.1037/h0093840

Peña, E. D., Bedore, L. M., and Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: two languages classify better than one. *Int. J. Lang. Commun. Disord.* 51, 192–202. doi: 10.1111/1460-6984.12199

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Congn.* 13, 253–262. doi: 10.1017/S1366728909990526

Saer, D. J. (1923). The effect of bilingualism on intelligence. *Br. J. Psychol.* 14, 25–38.

Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35, 606–621.

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., Gobet, F., et al. (2019). Near and far transfer in cognitive training: a second-order meta-analysis. *Collabra Psychol.* 5:18. doi: 10.1525/collabra.203

Sala, G., and Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Curr. Dir. Psychol. Sci.* 26, 515–520. doi: 10.1177/0963721417712760

Sala, G., and Gobet, F. (2019). Cognitive training does not enhance general cognition. *Trends Cogn. Sci.* 23, 9–20. doi: 10.1016/j.tics.2018.10.004

Salomon, G., and Perkins, D. N. (1989). Rocky roads to transfer: rethinking mechanism of a neglected phenomenon. *Educ. Psychol.* 24, 113–142. doi: 10.1207/s15326985ep2402_1

Siegelman, N., Bogaerts, L., Christiansen, M. H., and Frost, R. (2017b). Towards a theory of individual differences in statistical learning. *Philos. Transact. R. Soc. Biol. Sci.* 372:20160059. doi: 10.1098/rstb.2016.0059

Siegelman, N., Bogaerts, L., and Frost, R. (2017a). Measuring individual differences in statistical learning: current pitfalls and possible solutions. *Behav. Res. Methods* 49, 418–432. doi: 10.3758/s13428-016-0719-z

Simon, J. R., and Small, A. M. Jr. (1969). Processing auditory information: interference from an irrelevant cue. *J. Appl. Psychol.* 53, 433–435. doi: 10.1037/h0028034

Singh, L. (2021). Evidence for an early novelty orientation in bilingual learners. *Child Dev. Perspect.* 15, 110–116. doi: 10.1111/cdep.12407

Struys, E., Duyck, W., and Woumans, E. (2018). The role of cognitive development and strategic task tendencies in the bilingual advantage controversy. *Front. Psychol.* 9:1790. doi: 10.3389/fpsyg.2018.01790

Tversky, A. (1977). Features of similarity. *Psychol. Rev.* 84, 327–352. doi: 10.1037/0033-295X.84.4.327

Van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., et al. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behav. Sci.* 9:27. doi: 10.3390/bs9030027

Vīnerte, S., and Sabourin, L. (2019). Reviewing the bilingual cognitive control literature: can a brain-based approach resolve the debate? *Can. J. Exp. Psychol.* 73, 118–134. doi: 10.1037/cep0000174

Ware, A. T., Kirkovski, M., and Lum, J. A. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Front. Psychol.* 11:1458. doi: 10.3389/fpsyg.2020.01458

Weiss, D. J., Gerfen, C., and Mitchel, A. D. (2010). Colliding cues in word segmentation: the role of cue strength and general cognitive processes. *Lang. Cogn. Process.* 25, 402–422. doi: 10.1080/01690960903212254

Wiseheart, M., Viswanathan, M., and Bialystok, E. (2016). Flexibility in task switching by monolinguals and bilinguals. *Biling. Lang. Congn.* 19, 141–146. doi: 10.1017/S1366728914000273

Yoshioka, J. G. (1929). A study of bilingualism. *Pedagog. Semin. J. Genet. Psychol.* 36, 473–479. doi: 10.1080/08856559.1929.10532205